

RESULT 2

09v725 PRELIMINARY: PRT; 456 AA.

AC 09v725; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE CG11395 PROTEIN (GH14572p).

GN CG11395

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephyroidae; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

[1]

RP SEQUENCE FROM N.A.

RC MEDLINE=20196006; PubMed=10731132;

RX Adams M.D., Ceuliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H.C., Blaise R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Burtis K.C., Busan D.A., Butler C., Davenport L.B., Davies P.,

RA Cherry J.M., Cawley S., Dahlke C., Davemport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagha C.,

RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.R., Kechum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Mishina N.V., Modyarty C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissensbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA "The genome sequence of Drosophila melanogaster."

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN-Y, CN BW SP;

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,

RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,

RA Nunco J., Pacle J., Paragas V., Park S., Phooanavong S., Wan K.,

RA Yu C., Lewis S.E., Rubin G.M., Ceuliker S.,

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003804; AAF57889.1; -

DR EMBL; AY051489; AAK92913.1; -

DR FLYbase; FBgn0034200; CG11395.

SO SEQUENCE 456 AA; 48785 MW; D69162E132E6626F CRC64;

Query Match 12.6%; Score 78; DB 5; Length 456;
 Best Local Similarity 30.7%; Pred. No. 3.9;
 Matches 23; Conservative 11; Mismatches 37; Indels 4; Gaps 2;

OY 16 AOSFGRKCKPEQEVY-PGGGRSKRDPDLYQLQRFKSHSSLEGLKRLASQASNDPK-- 72
 DB 361 AOTIQIPHPHPOSSVYQIPGRRQSSNDVNDINLFTHTDPLSPPLSTAPCANRDPKSD 420
 OY 73 -ESTSPKRDMDFF 86
 DB 421 AEAAPESVYNORSLF 435

RESULT 3

022896 PRELIMINARY: PRT; 2364 AA.

AC 022896; 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 263.2 KDA PROTEIN.

GN C16D9.2.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI_TaxID=6239;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "None"

RL "Investigating biology. The C. elegans Sequencing Consortium."

RL Science 282:2012-2018(1998).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Gattung S., Le T.T.

RL "The sequence of C. elegans cosmid C16D9."

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Waterston R.

RT "Direct Submission."

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; U64858; AB18281.2; -

DR InterPro: IPR000719; Euk_Pkinase.

DR InterPro: IPR002290; Ser_thr_Pkinase.

DR InterPro: IPR001245; Tyr_Pkinase.

DR Pfam: PF00064; fn3.3.

DR Pfam: PF00069; Pkinase.2.

DR PRINTS; PR00109; TYRKINASE.

DR SMART; SM00060; FN3.3.

DR SMART; SM00220; S_TKc.1.

DR SMART; SM00219; TYRKc.1.

DR PROSITE; PSS0011; PROTEIN_KINASE_DOM.1.

KW ATP-binding; Hypothetical protein; Transferase.

SO SEQUENCE 2364 AA; 263180 MW; F2BC8C5E9E952B46 CRC64;

Query Match 12.5%; Score 77.5; DB 5; Length 2364;
 Best Local Similarity 38.6%; Pred. No. 35;
 Matches 22; Conservative 8; Mismatches 18; Indels 9; Gaps 3;

OY 65 SQASTDKESTSPKRDMDFFVGLMKRSVOPSDTVNQENVPSGILKYPRAE 121
 DB 2144 SODSTSSREPPSPSHR-MRDF-----IDTRDLRPPSPSHLNM-----SFGGEHYHEE 2191

RESULT 4

09SB51

```

ID 09SB51 PRELIMINARY; PRT; 1008 AA.
AC 09SB51;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HYPOTHEtical 110.6 KDA PROTEIN.
GN F22K18.240 OR ATG424560.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Berman M., Medler E., Wambutt R., Hohenisel J., Mewes H.W.,
RA Mayer K.F.X., Schueller C.
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Medler H., Medler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL033356; CAA23007.1; -.
DR EMBL; AL161561; CAB/9366.1; -.
DR InterPro; IPR001394; UCH-2.
DR InterPro; IPR002893; Znf-MYND.
DR Pfam; PF00442; UCH-1; 1.
DR Pfam; PF00443; UCH-2; 1.
DR Pfam; PF01753; ZF-MYND; 1.
DR PROSITE; PS50235; UCH_2_3; 1.
DR Hypothetical protein.
KW SEQUENCE 1008 AA; 110599 MW; F2DF3BD6E9039B9E CRC64;
SQ

Query Match 12.1%; Score 75; DB 10; Length 1008;
Best Local Similarity 33.0%; Pred. No. 22;
Matches 29; Conservative 8; Mismatches 41; Indels 10; Gaps 4;

QY 32 PGGRSKRDPDLYQLRFLK-----SHSLLEGILKALSOASTDPKSTSPKRDMDHF 86
DB - 896 PGGRSGNIOSEFYSSFORLQKILEDSASDSSSLFDSNDESCSCTDSTMD--DFADFI 953
QY 87 VG-LMGKRSVOPD--SPDVNQNENVPSPF 111
DB 954 FGDHQGRAGQSETPSPSTSSSSSSPPF 981

RESULT 5
Q9PFS8 PRELIMINARY; PRT; 1008 AA.
AC Q9PFS8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE UBIQUITIN-SPECIFIC PROTEASE 16.
GN UBP16.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-20567829; PubMed=11115897;
RA Van N., Doelling J.H., Falbel T.G., Durski A.M., Vierstra R.D.;
RT "The Ubiquitin-Specific Protease Family from Arabidopsis. AtUBP1 and 2

```

```

RT Are Required for the Resistance to the Amino Acid Analog Canavanine.";
PL Plant Physiol. 124:1828-1843(2000).
DR EMBL; AF302666; AAG42757.1; -.
DR MEROPS; C19; UPW; -.
DR InterPro; IPR001394; UCH-2.
DR InterPro; IPR002893; Znf-MYND.
DR Pfam; PF00442; UCH-1; 1.
DR Pfam; PF00443; UCH-2; 1.
DR Pfam; PF01753; ZF-MYND; 1.
DR PROSITE; PS50235; UCH_2_3; 1.
KW Protease.
SQ SEQUENCE 1008 AA; 110585 MW; 2BAC6F35ED506DFB CRC64;

Query Match 12.1%; Score 75; DB 10; Length 1008;
Best Local Similarity 33.0%; Pred. No. 22;
Matches 29; Conservative 8; Mismatches 41; Indels 10; Gaps 4;

QY 32 PGGRSKRDPDLYQLRFLK-----SHSLLEGILKALSOASTDPKSTSPKRDMDHF 86
DB 896 PGGRSGNIOSEFYSSFORLQKILEDSASDSSSLFDSNDESCSCTDSTMD--DFADFI 953
QY 87 VG-LMGKRSVOPD--SPDVNQNENVPSPF 111
DB 954 FGDHQGRAGQSETPSPSTSSSSSSPPF 981

RESULT 6
Q9XYV1 PRELIMINARY; PRT; 301 AA.
AC Q9XYV1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CYCLIN-DEPENDENT PROTEIN KINASE CDK2.
GN CDK2.
OS Parametium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;
OC Parametium.
OX NCBI_TaxID=5888;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=515;
RC MEDLINE=99446661; PubMed=10519216;
RA Zhang H., Berger J.D.;
RT "A novel member of the cyclin-dependent kinase family in Parametium
RT tetraurelia."
RL J. Eukaryot. Microbiol. 46:482-491(1999).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF126147; AAD34354.1; -.
DR HSP; F24941; HCL.
DR InterPro; IPR00719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 301 AA; 34675 MW; E839F1A5EA05CB5 CRC64;

Query Match 12.0%; Score 74.5; DB 5; Length 301;
Best Local Similarity 29.3%; Pred. No. 5.6;
Matches 34; Conservative 10; Mismatches 45; Indels 27; Gaps 5;

QY 8 TATLAPSLQSGFACKEPOEEVVP-----GGGRSKRDP-DLYOL-----LQ 48
DB 146 TQADFGILARAGLPKITYTHETITMYRAPELLGGRYSIPVIMSLGCTFAEAGKR 205
QY 49 RLKSHSLLEGILKALSOASTDPKSTSPKRDMDHFVGLMGKRSVOPDSTPDVN 104
DB 206 PLFCGSEIDQLFKIRKIMGF--PKESTWPGVSTLPDF-----KSTFRMPTPTN 253

```

```

RESULT 7
ID 061048 PRELIMINARY; PRT; 376 AA.
AC 061048;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FORMIN BINDING PROTEIN 21 (FORMIN BINDING PROTEIN FBP 21).
GN WBP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=98393691; PubMed=9724750;
RA Bedford M.T., Reed R., Leder P.;
RT "WW domain-mediated interactions reveal a spliceosome-associated
RT protein that binds a third class of proline-rich motif: the proline
RT glycine and methionine-rich motif";
RL Proc. Natl. Acad. Sci. U.S.A. 95:10602-10607(1998).
RN 12
RP SEQUENCE OF 129-195 FROM N.A.
RX STRAIN=EV8;
RC MEDLINE=96183189; PubMed=8605874;
RA Chan D.C., Bedford M.T., Leder P.;
RT "Formin binding proteins bear WW/PW domains that bind proline-rich
RT peptides and functionally resemble SH3 domains.";
RL EMBL J. 15:1045-1054(1996).
DR EMBL; AF071184; AAC34810.1; -.
DR MGI; 040746; AAC52474.1; -.
DR MGI; 109568; Wbp4.
DR InterPro: IPR001202; WW.
DR InterPro: IPR000690; ZnF_maf1rn.
DR InterPro: IPR003604; ZnF_U1.
DR Pfam: PR00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR SMART; SM00451; ZnF_U1; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; UNKNOWN_1.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
SQ SEQUENCE 376 AA; 42042 MW; 8ED9334CC8CD6D04 CRC64;

Query Match 12.0%; Score 74.5; DB 11; Length 376;
Best Local Similarity 28.9%; Pred. No. 7.4;
Matches 28; Conservative 16; Mismatches 32; Indels 21; Gaps 5;

QY 25 EPOEEVPPGG---RSKRNPDIYQLLO--RLFKSHSSLEGLKALSOASTDPK----- 72
DB 193 EKPEDFIPHGCVLSSKDSGLPTLEAKSSDSHSDSEGEQKKAGASTETKLLIRFK 252
QY 73 -ESTSPEKRDMDHFFVG--LMGKRSVOPDSPDVNOE 106
DB 253 EKNKSTKTR-----IGPEIQEKSTPNQNSNTNTE 283

RESULT 8
ID 09SBK6 PRELIMINARY; PRT; 392 AA.
AC 09SBK6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FLORAL NECTARY-SPECIFIC PROTEIN.
GN NTR1.
OS Brassica rapa subsp. pekinensis (Chinese cabbage) (celery cabbage).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC eucosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=51351;
RN 11

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=20267007; PubMed=10809010;
RA Song J.T., Seo H.S., Song S.I., Lee J.S., Choi Y.D.;
RT "NTR1 encodes a floral nectary-specific gene in Brassica campestris L.
RT ssp. pekinensis.";
RL Plant Mol. Biol. 42:647-655(2000).
DR EMBL; AF179222; AAF22289.1; -.
SQ SEQUENCE 392 AA; 43815 MW; 25H78530E93B5757 CRC64;

Query Match 12.0%; Score 74; DB 10; Length 392;
Best Local Similarity 26.3%; Pred. No. 8.8;
Matches 35; Conservative 17; Mismatches 51; Indels 30; Gaps 7;

QY 10 ILAFLAOSFGAVCKEPOEEVPPG-----GRSKRPD-----LYQLQLRFLK 52
DB 207 VVALQFOQDFSVFLKRSSEELVPGGRWVLSFLGRSSPDPTRESCYQELLALMSLAK 266
QY 53 SHSSLEGLKALS-QASIDPKKSTSPKRDMDHFFVGLMKRSVQ--PDSPTDVNOENV- 108
DB 267 -----EGTIEENIDAFNAPVYMAASPEIKM---AIKEGGSFSDRLRLEISPDWEGSGIS 318
QY 109 -PSFGILKYPRA 120
DB 319 DSDYDIVRKPRA 331

RESULT 9
ID 09FF55 PRELIMINARY; PRT; 597 AA.
AC 09FF55;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PROTEIN DISULPHIDE ISOMERASE-LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=COLUMBIA;
RC MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.,
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT pl clones.";
RL DNA Res. 4:215-230(1997).
DR EMBL; AB005246; BAB09837.1; -.
DR HSP; P07237; IMEK.
DR InterPro: IPR000886; ER_target.
DR InterPro: IPR000063; ThioRed.
DR Pfam: PF00085; ThioRed; 3.
DR PRINTS; PR00421; THIOREDOXIN.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Isomerase.
SQ SEQUENCE 597 AA; 66357 MW; 5A8FC7E72AA64B2B CRC64;

Query Match 11.9%; Score 73.5; DB 10; Length 597;
Best Local Similarity 24.6%; Pred. No. 17;
Matches 33; Conservative 21; Mismatches 35; Indels 45; Gaps 7;

QY 2 RIMLTF--TAILAFS--LAOSFGA-----VCKEPOEEVPPG----- 34
DB 4 RVLLEFSLTLLILFSAVSPFAASSDDVDDEDLSFLEDKEDVDVPGADSLSSSTGFDF 63
QY 35 -GRSKRNPDIYQLQLRFLKSHSSLEGLKALSQASIDPKKSTSPKRDMDHFFVGLMKR 93
DB 64 EGGEEDPDMT-----NDDDEGDFSDLGPNPDDPLPTPEIDKD-----VVVIKER 111

```



```

OY 94 SVQPSPIVDVNOEN 107
    111 : 1
Db 112 NF-----TDVIENN 120

RESULT 10
O9D4S8 PRELIMINARY; PRT; 172 AA.
ID O9D4S8;
AC O9D4S8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 4930563P03RIK PROTEIN.
GN 4930563P03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinozawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Felschmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Strubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Bottelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gusting S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK016213; BAB30150.1;
DR MGI; MGI:1914905; 4930563P03RIK.
SQ SEQUENCE 172 AA; 19043 MW; E3243D9122FC5811 CRC64;

Query Match 11.6%; Score 72; DB 11; Length 172;
Best Local Similarity 26.4%; Pred. No. 5.2;
Matches 32; Conservative 16; Mismatches 55; Indels 18; Gaps 6;

OY 2 R1MLFTALFSLAQSGAVCKEPOEEVPGGGRKRDPPDYOLLQRLFKSHSLGGL 61
    111 : 111
Db 3 R1LVVAVANQPHQ-QGMFEELPESGDGQPGPARRRKQPSSEAMPYLTKEDLESDM 61

OY 62 K-----ALSCASTDPK-ESTSPKRDMDFFVGLMGKRSYQDPS--TDVNOE---NVPSF 111
    111 : 111
Db 62 KEVDILGSGSDSDIEKKRPEDQNEQ-----ERAPKPRKPRAPGIRRQPVLPSS 114

OY 112 G 112
Db 115 G 115

RESULT 11
O9RZ15 PRELIMINARY; PRT; 969 AA.
ID O9RZ15;
AC O9RZ15;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MCRB-RELATED PROTEIN.
GN DRB0143.

```

```

OS Deinococcus radiodurans.
OG Plasmid Mpl.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036696; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uitterlidge T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans RI.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001826; AAF12568.1;
DR TIGR; DRB0143;
KW Plasmid; Complete proteome.
SQ SEQUENCE 969 AA; 105295 MW; 3076287B1D0994A4 CRC64;

Query Match 11.6%; Score 72; DB 16; Length 969;
Best Local Similarity 26.4%; Pred. No. 45;
Matches 33; Conservative 14; Mismatches 54; Indels 24; Gaps 4;

OY 7 FPAIIAFLSLAQSGAVCKEPOEEVPGGGRKRDPPDYOLLQRLFKSHSLGGL 53
    111 : 111
Db 368 FGAFLFLPLPGPREAQAQLOQYMSGVRSLRGLSLPEHLSDKFPFLDANLTY 427

OY 54 HSLLEGLKALSOASTDPKSTSPKRDMDFFVGLMGKRSYQDPSPTVNOENVPSCFI 113
    111 : 111
Db 428 LDALVDLDLGLMR-----PEESTSE-----IGVGRPVAPPAARVPPGVPPLNGI 476

OY 114 LKYP 118
Db 477 LVGP 481

RESULT 12
O9FT99 PRELIMINARY; PRT; 208 AA.
ID O9FT99;
AC O9FT99;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 23.7 KDA PROTEIN.
F8LI5_90.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Terry N., Ardiles W., Buyschaert C., Daseville R.,
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA Villarejo R., Gléhen J., Van Montagu M., Bancroft I., Meves H.W.,
RA Ruid S., Lemcke K., Mayer K.F.X.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL392174; CAC08336.1;
KW Hypothetical protein.
SQ SEQUENCE 208 AA; 23740 MW; F77DBBF0B4AEF81B CRC64;

Query Match 11.6%; Score 71.5; DB 10; Length 208;
Best Local Similarity 26.7%; Pred. No. 7.5;

```

Matches 28; Conservative 20; Mismatches 48; Indels 9; Gaps 4;

OY 14 SLAOSGAVCK--EPQEEVVGCGSKRDPDYOLLOLRLFKSHSLLEGILRA-LSQAST 69
 DB 16 SSRSGEGECHLFFEFYNEVSVGDLRL---PLTMVEELAKHHGLNLTSLTSDLSNSW 71
 OY 70 DPEKSTSPKRDHDFEYGLMGKRSVOPDPTDVNOENVPSEGLIKY 114
 DB 72 SP-ATGIPSRQTLNOYELTYHSLTPVTPETPKGTEVELPARGVL 115

RESULT 13

OY919V9 PRELIMINARY; PRT; 234 AA.
 AC OY919V9;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE GAG POLYPROTEIN (FRAGMENT).
 GN GAG.
 OS Dendragapus obscurus (blue grouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Tetraonidae; Dendragapus.
 OX NCBI_TaxID=90755;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20219390; PubMed=10756010;
 RA Dimcheff D.E., Drovetski S.V., Krishnan M., Mindell D.P.;
 RT "Cospeciation and horizontal transmission of avian sarcoma and
 leukosis virus gag genes in galliform birds."
 RT J. Virol. 74:3984-3995(2000).
 DR EMBL; AF225347; AAF64722.1;
 DR HSSP; P03322; 1A6S.
 DR InterPro; IPR004028; Retro_M.
 DR Pfam; PF02813; Retro_M.1.
 KM Polyprotein.
 KN NON_TER
 FT NON_TER 1 234
 FT SEQUENCE 234 AA; 24333 MW; 61ADCC04CBCF4C9F4 CRC64;
 SQ

Query Match 11.6%; Score 71.5; DB 13; Length 234;
 Best Local Similarity 32.7%; Pred. No. 8.7;
 Matches 36; Conservative 12; Mismatches 45; Indels 17; Gaps 7;

OY 14 SLAOSGAVCKEPOEEVVGCGSKRDPDYOLLOLRLFKSHSLLE---GL-LKALSQAS 68
 DB 21 SLQKREG-LTSPSDIYSGSW----DPTVNLISQVLMELGKSEELRTGVLKALKAKAR 75
 OY 69 TTPKESTSPKRDHDFEYGLMGKRSVOPDPTDVNOENVPSEGLIKYPP.118
 DB 76 ---EEQVPEQA---KFLGLGGGK-VSPGPTDEKEKMAAQETKVAP 118

RESULT 14

OY91T08 PRELIMINARY; PRT; 495 AA.
 AC OY91T08;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE GBAIAF00675.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20277480; PubMed=10819329;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 features of the regions of 4,504,864 bp covered by sixty pl and YAC
 clones."
 RT DNA Res. 7:131-135(2000).
 RL EMBL; AB024036; BAB02816.1;
 DR EMBL; AB024036; BAB02816.1;
 SQ SEQUENCE 495 AA; 56812 MW; 19900B9B9B9F9C6E CRC64;

Query Match 11.6%; Score 71.5; DB 10; Length 495;
 Best Local Similarity 27.4%; Pred. No. 22;
 Matches 29; Conservative 16; Mismatches 44; Indels 17; Gaps 4;

OY 14 SLAOSGAVCKEPOEEVVGCGSKRDPDYOLLOLRLFKSHSLLEGILKALSQASTDPKE 73
 DB 373 TLREIKVCSYPRRQ-----SKRLPSMAKL-----WSLLEGMGKRSVNSARKSTVE 418
 OY 74 STSPKRDHDFEYGLMGKRSVOPDPTDVNOENVPSEGLIKY 117
 DB 419 MVSPTGSKNGGFNTLDVGWSSSPDS-AMANLNGKRGKCIEMP 463

RESULT 15

OY9VZL5 PRELIMINARY; PRT; 463 AA.
 AC OY9VZL5;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE CG14980 PROTEIN.
 GN CG14980.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadien E., Caceres A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy C., Morris J., Moshrefi A.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheelke F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski W.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

THE
FEDERAL
BUREAU OF INVESTIGATION
UNITED STATES DEPARTMENT OF JUSTICE
WASHINGTON, D. C. 20535

TO : DIRECTOR, FBI (100-442100) FROM : SAC, NEW YORK (100-100000) (P)
SUBJECT: [REDACTED] (C) (U) (P)
RE: [REDACTED] (C) (U) (P)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.,

OM protein - protein search, using sw model

Run on: May 24, 2002, 17:03:36 ; Search time 11.87 Seconds

(without alignments)
394,698 Million cell updates/sec

Title: US-09-852-659-68

Perfect score: 619
Sequence: 1 MRIMLFTAILAFSLAQSFQ.....DVQENVPSFGILIKYPPRAE 121

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	619	100.0	121	1	TKNK_HUMAN
2	416.5	67.3	126	1	TKNK_BOVIN
3	373.5	60.3	116	1	TKNK_MOUSE
4	353.5	57.1	116	1	TKNK_RAT
5	77	12.4	221	1	FA5_BOVIN
6	72	11.6	921	1	AKH_DAUCA
7	70.5	11.4	237	1	HK31_MOUSE
8	70	11.3	343	1	MAT2_COCHF
9	69	11.1	356	1	E2F3_MOUSE
10	69	11.1	465	1	E2F3_HUMAN
11	69	11.1	2096	1	BP28_DROME
12	68	11.0	130	1	TKN1_MESAU
13	68	11.0	302	1	RDGC_XYLEFA
14	67.5	10.9	355	1	CKX5_RAT
15	67.5	10.9	1374	1	WCAP_HSV11
16	67.5	10.9	2464	1	MAFB_MOUSE
17	67	10.8	513	1	PIX1_DROME
18	67	10.8	1028	1	OVO_DROME
19	66.5	10.7	978	1	SIR3_YEAST
20	66.5	10.7	2353	1	CCAH_HUMAN
21	66	10.7	130	1	TKN1_RAT
22	66	10.7	293	1	KRP_CHAV
23	66	10.7	452	1	E2BG_RAT
24	66	10.7	823	1	YRR1_CABEL
25	65.5	10.6	129	1	TKN1_HUMAN
26	65.5	10.6	130	1	TKN1_BOVIN
27	65.5	10.6	450	1	NCAP_RABVP
28	65.5	10.6	910	1	DNM1_MYCPN
29	65.5	10.6	920	1	TMU1_YEAST
30	65	10.5	168	1	ANPL_CABEL
31	65	10.5	282	1	PE11_YEAST
32	65	10.5	927	1	BZAT_RAT
33	65	10.5	1087	1	POM1_SCHPO

34	65	10.5	1270	1	DDX9_HUMAN	Q08211	homo sapien
35	64.5	10.4	130	1	TKN1_MOUSE	P41539	mus musculus
36	64.5	10.4	386	1	CVCB_PEA	P13919	Pisum sativ
37	64.5	10.4	991	1	Y419_HUMAN	O43303	homo sapien
38	64.5	10.4	1091	1	MSH3_MOUSE	P13703	mus musculus
39	64.5	10.4	3119	1	HD_MOUSE	P42859	mus musculus
40	64	10.3	201	1	SDC2_RAT	P34900	rattus norv
41	64	10.3	307	1	CIW7_HUMAN	Q9Y2U2	homo sapien
42	64	10.3	920	1	AKH1_MAIZE	P49079	zea mays (m
43	64	10.3	3148	1	HD_FUGRU	P51112	fugu rubrip
44	63.5	10.3	281	1	EPID_MOUSE	P57776	mus musculus
45	63.5	10.3	469	1	MM13_XENLA	Q10835	xenopus lae

ALIGNMENTS

RESULT 1
TKNK_HUMAN STANDARD; PRT; 121 AA.
ID TKNK_HUMAN
AC Q9UHF0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurokinin B precursor (NRB) (Neupomedin K) (ZNEUROK1).
GN TAC3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sheppard P., Jelinek L., Whitmore T., Blumberg H., Lehner J.,
RL O'Hara P.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=20322570; PubMed=10866201;
RA Page N.M., Woods R.J., Gardiner S.M., Lomthasong K., Gladwell R.T.,
RL Bullin D.J., Manyonda I.T., Lowry P.J.;
RT "Excessive placental neurokinin B secretion during the third trimester
causes pre-eclampsia."
RL Nature 405:797-800(2000).
CC -I- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
MUSCLES (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- DEVELOPMENTAL STAGE: In pregnancy, the expression of NRB is
confined to the outer syncytiotrophoblast of the placenta,
significant concentrations of NRB can be detected in plasma as
early as week 9, and plasma concentrations of NRB are grossly
elevated in pregnancy-induced hypertension and pre-eclampsia.
CC -I- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF186112; AAF01430.1; -
CC EMBL: AF216586; AAF76980.1; -
CC MIM: 162330; -
CC InterPro: IPR003635; Neurokinin.
CC InterPro: IPR002040; Tachykinin.
CC PRODOM: PD020370; Neurokinin; 1.
CC PROSITE: PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
KW Amidation; Signal.

FT SIGNAL 1 16 POTENTIAL.
 FT PROPEP 17 78 BY SIMILARITY.
 FT PEPTIDE 81 90 NEUROKININ B.
 FT PROPEP 94 121 BY SIMILARITY.
 FT MOD_RES 90 90 AMIDATION (G-91 PROVIDE AMIDE GROUP) (BY
 FT SEQUENCE 121 AA; 13438 MW; 14C9AFE2EE9E9DECA CRC64;

Query Match 100.0%; Score 619; DB 1; Length 121;
 Best Local Similarity 100.0%; Pred. No. 1.3e-52;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRIMLFTAILAFSLAOSFGAVCKEPEEYVPGGSRKRDPLXOLRLFKSHSLEGL 60
 DB 1 MRIMLFTAILAFSLAOSFGAVCKEPEEYVPGGSRKRDPLXOLRLFKSHSLEGL 60
 OY 61 LKALSQASTDPKESTSPKRDMDFFVGLMKRSYQPDSPDVNOENVPSEGLKYPRA 120
 DB 61 LKALSQASTDPKESTSPKRDMDFFVGLMKRSYQPDSPDVNOENVPSEGLKYPRA 120
 OY 121 E 121
 DB 121 E 121

RESULT 2

TKNK_BOVIN

ID_TKNK_BOVIN

STANDARD: PRT: 126 AA.

AC P0858;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurokinin B precursor (NRB) (Neuromedin K).
 GN TAC3 OR NKNB
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCB1_TaxID=9913;
 RN [1]

RP SEQUENCE FROM N.A. PubMed-3462746;
 RA MEDLINE-86313713; Nawa H., Nakanishi S.;
 RA Kotoh H., Hoshimaru M., Nawa H., Nakanishi S.;
 RT "Structure and gene organization of bovine neuromedin K precursor";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:7074-7078(1986).
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M14351; AAA30723.1; -;
 DR EMBL; M14347; AAA30723.1; JOINED.
 DR EMBL; M14348; AAA30723.1; JOINED.
 DR EMBL; M14349; AAA30723.1; JOINED.
 DR EMBL; M14350; AAA30723.1; JOINED.
 DR PIR; A25905; A25905.
 DR InterPro; IPR003635; Neurokinin.
 DR InterPro; IPR002040; Tachykinin.
 DR ProDom; PD020370; Neurokinin.1.
 DR PROSITE; PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
 KW Amidation; Signal.

FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 83
 FT PEPTIDE 86 95 NEUROKININ B.
 FT PROPEP 99 126
 FT MOD_RES 95 95 AMIDATION (G-96 PROVIDE AMIDE GROUP).
 FT SEQUENCE 126 AA; 13871 MW; 446EF433498EC059 CRC64;

Query Match 67.3%; Score 416.5; DB 1; Length 126;
 Best Local Similarity 65.9%; Pred. No. 3.3e-33;
 Matches 83; Conservative 15; Mismatches 23; Indels 5; Gaps 2;

OY 1 MRIMLFTAILAFSLAOSFGAVCKEPEEYVPGGSRKRDPLXOLRLFKSH-S 55
 DB 1 MRSTLFTAILAFSLAOSFGAVCKEPEEYVPGGSRKRDPLXOLRLFKSH-S 60
 OY 56 SLEGLLKALSOASTDPKESTSPKRDMDFFVGLMKRSYQPDSPDVNOENVPSEGLK 115
 DB 61 SLDELKMLSKASVGPKEPSLPQKRDMHDFVGLMGKRNLPDPEVDINQENIPSEGTFK 120
 OY 116 YPPAE 121
 DB 121 YPPAE 126

RESULT 3

TKNK_MOUSE

ID_TKNK_MOUSE

STANDARD: PRT: 116 AA.

AC P35099;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurokinin B precursor (NRB) (Preprotachykinin B) (PPT-
 DE B).
 GN TAC3 OR NKNB OR TAC2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCB1_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RA STRAIN=ICR; TISSUE=Brain;
 RA Kato K., Muneoka E., Hosaka M., Murakami K., Nakayama K.;
 RT "Cloning and sequence analysis of mouse cDNAs encoding
 RT preprotachykinin A and B";
 RL Biomed. Res. 14:253-259(1993).

CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; D14423; BAA03316.1; -;
 DR MGD; MGI:98476; Tac2.
 DR InterPro; IPR003635; Neurokinin.
 DR InterPro; IPR002040; Tachykinin.
 DR ProDom; PD020370; Neurokinin.1.
 DR PROSITE; PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
 KW Amidation; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 79
 FT PEPTIDE 82 91 NEUROKININ B.
 FT PROPEP 95 116

FT MOD_RES 91 91 AMIDATION (G-92 PROVIDE AMIDE GROUP).
SQ SEQUENCE 116 AA: 12809 MW: BF6E89373E2031CC CRC64;

Query Match 60.3%; Score 373.5; DB 1; Length 116;
Best Local Similarity 65.0%; Pred. No. 3.9e-29;
Matches 78; Conservative 15; Mismatches 18; Indels 9; Gaps 3;

OY 1 MRIMLFTALILAFSLAOSFGAVCKEPOEVEVPGGSRKRDPLYO-----LLQRLFKSHS- 55
DB 1 MMSAMLFAAVALLSLAMTFGAVCEPQEO-----GGRLSKDLSLLPPLRLRLDLSRI 56
OY 56 SLEGLLKALISQASTDPKSTSPKRDMDHDFYGLMGKRSVQPSDPDVOENVPSEGLIK 115
DB 57 SLEGLLKALISQASTDPKSTSPKRDMDHDFYGLMGKRSVQPSDPDVOENVPSEGLIK 116

RESULT 4

TRNK_RAT 4 STANDARD: PRT: 116 AA.

AC P08435; 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurokinin B precursor (NKB) (Neuromedin K).
GN TAC3 OR NKB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

RP SEQUENCE FROM N.A.
RA MEDLINE=88051833; PubMed=3479225;
RA Bonner T.I., Afolter H.-U., Young A.C., Young W.S. III;
RT "A CDNA encoding the precursor of the rat neuropeptide, neurokinin B."
RT B."

RL Brain Res. 388:243-249(1987).

CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOLVE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC (SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M16410; AAA41711.1; -

DR PIR: A43779; A43779.

DR InterPro: IPR003635; Neurokinin.

DR InterPro: IPR002040; Tachykinin.

DR ProDom: PD020370; Neurokinin; 1.

DR PROSITE: PS00267; TACHYKININ; 1.

KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;

KW Amidation; Signal.

FT SIGNAL 1 20 POTENTIAL.

FT PROPEP 21 79 NEUROKININ B.

FT PEPTIDE 82 91

FT PROPEP 95 116

FT MOD_RES 91 91 AMIDATION (G-92 PROVIDE AMIDE GROUP).

SQ SEQUENCE 116 AA: 12659 MW: C73BC67E2BAF88C CRC64;

Query Match 57.1%; Score 353.5; DB 1; Length 116;
Best Local Similarity 61.7%; Pred. No. 3.2e-27;
Matches 74; Conservative 16; Mismatches 21; Indels 9; Gaps 3;

OY 1 MRIMLFTALILAFSLAOSFGAVCKEPOEVEVPGGSRKRDPLYO-----QLQLRFLFSHS- 55

DB 1 MMSAMLFAAVALLSLAMTFGAVCEPQEO-----GGRLSKDLSLLPPLRLRLDLSRI 56
OY 56 SLEGLLKALISQASTDPKSTSPKRDMDHDFYGLMGKRSVQPSDPDVOENVPSEGLIK 115
DB 57 SLEGLLKALISQASTDPKSTSPKRDMDHDFYGLMGKRSVQPSDPDVOENVPSEGLIK 116

RESULT 5

FA5_BOVIN 5 STANDARD: PRT: 2211 AA.

ID FA5_BOVIN 028107; 028108;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Coagulation factor V precursor (Activated protein C cofactor).

GN F5.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

GN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=92147638; PubMed=1737753;

RA Guinlo E.R., Esmen C.T., Mann K.G., Macgillivray R.T.;

RT "The complete cDNA sequence of bovine coagulation factor V."

RL J. Biol. Chem. 267:2971-2978(1992).

CC -1- FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES

CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.

CC -1- SUBUNIT: Factor Va is composed of a heavy chain and a light

CC chain, noncovalently bound. The interaction between the two chains

CC is calcium-dependent.

CC -1- DOMAIN: Domain B contains 29.5 x 9 AA tandem repeats, and 2 x 14

CC AA repeats.

CC -1- PFM: Thrombin activates factor V proteolytically to the active

CC cofactor, factor Va (formation of a heavy chain at the N-

CC terminus and a light chain at the C-terminus).

CC -1- PFM: SULFATION IS REQUIRED FOR EFFICIENT THROMBIN CLEAVAGE AND

CC ACTIVATION AND FOR FULL PROCOAGULANT ACTIVITY (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 3 F5/F8 TYPE A DOMAINS; EACH IS COMPOSED OF

CC 2 PLASTOCYANIN-LIKE REPEATS.

CC -1- SIMILARITY: CONTAINS 2 F5/F8 TYPE C DOMAINS.

CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.

CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M81440; AAA30512.1; -

DR EMBL: M81441; AAA30513.1; -

DR HSP: P00450; IKCW.

DR InterPro: IPR001117; Cu-oxidase.

DR InterPro: IPR000421; FA58_C.

DR Pfam: PF00394; Cu-oxidase; 3.

DR Pfam: PF00754; F5_F8_type_C; 2.

DR SMART: SM00231; FA58C; 2.

DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.

DR PROSITE: PS01285; FA58C_1; 2.

DR PROSITE: PS01286; FA58C_2; 2.

KW Blood coagulation; Glycoprotein; Sulfation; Calcium; Signal; Zymogen;

KW Repeat.

FT SIGNAL 1 28 POTENTIAL.

FT CHAIN 29 2211 COAGULATION FACTOR V.

FT CHAIN 29 2211 HEAVY CHAIN (BY SIMILARITY).

FT PEPTIDE 742 1564 ACTIVATION PEPTIDE (CONNECTING REGION)

FT PEPTIDE 742 1564 (BY SIMILARITY).

FT CHAIN 1565 2211 LIGHT CHAIN (BY SIMILARITY).

```

FT DOMAIN 30 327 F5/8 TYPE A 1.
FT DOMAIN 30 193 PLASTOCYANIN-LIKE 1.
FT DOMAIN 203 327 PLASTOCYANIN-LIKE 2.
FT DOMAIN 348 666 F5/8 TYPE A 2.
FT DOMAIN 348 525 PLASTOCYANIN-LIKE 3.
FT DOMAIN 535 686 PLASTOCYANIN-LIKE 4.
FT SIMILAR 899 915 TO 17 AA REPEATS IN HUMAN FAS.
FT DOMAIN 1124 1151 2 x 14 AA TANDER REPEATS.
FT REPEAT 1138 1151 1-1.
FT DOMAIN 1188 1453 1-2.
FT REPEAT 1188 1196 30 x 9 AA APPROXIMATE TANDER REPEATS OF
FT REPEAT 1197 1205 [AS]-L-S-P-D-[LP]-[GS]-O-[TE].
FT REPEAT 1206 1214 2-1.
FT REPEAT 1223 1223 2-2.
FT REPEAT 1224 1232 2-3.
FT REPEAT 1233 1241 2-4.
FT REPEAT 1242 1250 2-5.
FT REPEAT 1251 1259 2-6.
FT REPEAT 1260 1268 2-7.
FT REPEAT 1277 1277 2-8.
FT REPEAT 1278 1286 2-9.
FT REPEAT 1287 1295 2-10.
FT REPEAT 1296 1304 2-11.
FT REPEAT 1305 1313 2-12.
FT REPEAT 1314 1332 2-13.
FT REPEAT 1331 1331 2-14.
FT REPEAT 1332 1331 2-15.
FT REPEAT 1333 1331 2-16.
FT REPEAT 1332 1340 2-17.
FT REPEAT 1341 1349 2-18.
FT REPEAT 1350 1358 2-19.
FT REPEAT 1367 1367 2-20.
FT REPEAT 1368 1376 2-21.
FT REPEAT 1377 1385 2-22.
FT REPEAT 1386 1394 2-23.
FT REPEAT 1395 1403 2-24.
FT REPEAT 1404 1412 2-25.
FT REPEAT 1413 1421 2-26.
FT REPEAT 1422 1430 2-27.
FT REPEAT 1431 1439 2-28.
FT REPEAT 1440 1444 2-29 (PARTIAL).
FT REPEAT 1445 1453 2-30.
FT DOMAIN 1369 1890 F5/8 TYPE A 3.
FT DOMAIN 1569 1738 PLASTOCYANIN-LIKE 5.
FT DOMAIN 1748 1890 PLASTOCYANIN-LIKE 6.
FT DOMAIN 1894 2048 F5/8 TYPE C 1.
FT DOMAIN 2053 2208 F5/8 TYPE C 2.
FT SITE 741 742 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1034 1035 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1564 1565 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DISULFID 167 193 PROBABLE.
FT DISULFID 499 525 PROBABLE.
FT DISULFID 1712 1738 PROBABLE.
FT DISULFID 1894 2048 PROBABLE.
FT DISULFID 2053 2208 BY SIMILARITY.
FT MOD_RES 697 697 SULFATION (POTENTIAL).
FT MOD_RES 701 701 SULFATION (POTENTIAL).
FT MOD_RES 730 730 SULFATION (POTENTIAL).
FT MOD_RES 1513 1513 SULFATION (POTENTIAL).
FT MOD_RES 1529 1529 SULFATION (POTENTIAL).
FT MOD_RES 1537 1537 SULFATION (POTENTIAL).
FT MOD_RES 1541 1541 SULFATION (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 756 756 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 780 780 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 902 902 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 952 952 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1053 1053 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1062 1062 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1094 1094 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1451 1451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1490 1490 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1550 1550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1690 1690 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1839 1839 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1997 1997 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2196 2196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 587 592 NETLPA -> T (IN VARIANT 2).
SQ SEQUENCE 2211 AA; 248981 MM; CBBF90B738667C45 CRC64;

Query Match 12.4%; Score 77; DB 1; Length 2211;
Best Local Similarity 28.6%; Pred. No. 25;
Matches 32; Conservative 13; Mismatches 39; Indels 28; Gaps 6;

QY 14 SLAOSFGAVCKRP---QEEVNVGGGRSKRDPDLYQ-----LQRLFKSHSLEGLAL 64
Db 1269 SLSPDLGGTALSPDPGQESLSPDLGGTALSPDLGGTALSPDLGGTALSPDLGGTALSPDL 1328
QY 65 SQASTDP--KSTSPKRDMDHFEVGLMGKSKVQPSDPTDYNQBNV-PSFG 112
Db 1329 GGTALSPDPGQESLSPD-----LGGTSLSP-----DLGGTSLSPDLG 1365

RESULT 6
AKH_DAUCA STANDARD: PRT; 921 AA.
AC P37142:
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional aspartokinase/homoserine dehydrogenase, chloroplast
DE precursor (AK-HD) (AK-HSDH) [Includes: Aspartokinase (EC 2.7.2.4);
DE Homoserine dehydrogenase (EC 1.1.1.3)] (Fragment).
OS Daucus carota (carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93283634; PubMed=8507831.
RA Weisemann J.M., Matthews B.F.;
RT "Identification and expression of a cDNA from Daucus carota encoding
RT a bifunctional aspartokinase-homoserine dehydrogenase.";
RL Plant Mol. Biol. 22:301-312(1993).
CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P)(+) = L-aspartate 4-
CC semialdehyde + NAD(P)H.
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
CC aspartate.
CC -1- PATHWAY: CATALYZES 2 NONCONSECUTIVE REACTIONS IN THE COMMON
CC BIOSYNTHETIC PATHWAY LEADING FROM ASP TO DIAMINOPIMELATE AND LYS,
CC TO MET, AND TO THR AND ILE.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ASPARTOKINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC HOMOSERINE DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L11529; AAL16972.1; -
DR PIR: S35160; S35160.
DR InterPro: IPR002912; ACT.
DR InterPro: IPR001048; Aspartokinase.
DR InterPro: IPR001341; Aspartokinase.
DR InterPro: IPR001342; Homoserine_dh.
DR Pfam: PF00696; aaknase; 1.
DR Pfam: PF01842; ACT; 2.
DR Pfam: PF00742; Homoserine_dh; 1.
DR PROSITE: PS00324; ASPARTOKINASE; 1.
DR PROSITE: PS01042; HOMOSER_DHEXENASE; 1.
DR Transferrase; Kinase; Oxidoreductase; Methionine biosynthesis; NADP;
KW Multifunctional enzyme; Chloroplast; Transit peptide.
FT NON_TER 1 1
FT TRANSIT 1 87
FT CHAIN 88 921
FT DOMAIN 88 ?
FT DOMAIN ? ?
FT DOMAIN ? 921
FT NP_BIND ? ?
SQ SEQUENCE 921 AA: 100226 MW: 90890392DA70A996 CRC64;

Query Match 11.6%; Score 72; DB 1; Length 921;
Best Local Similarity 31.2%; Pred. No. 28;
Matches 29; Conservative 13; Mismatches 41; Indels 10; Gaps 3;

OY 30 VYGGGRSRDDLYQLQLRFRSHSLF-GLIKALISQASTPKSTSEKRDMDHFEVG 88
DB 243 VYVAGSNVDPDYSESRLEKWFSSNOCITVATGFATSPQNIPTLRKDGSDFSAA 302
OY 89 LMGK--RSYQSPSPMDVNGENPSPGILKYP 119
DB 303 LMGALLRAGQVITWDVN-----GVYSADPR 328

RESULT 7
ID HK31_MOUSE STANDARD; PRT: 237 AA.
AC P97436; 009087;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE -Homeobox protein NKX-3.1.
GN NKX3A OR NKX3-1 OR NKX-3.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA MEDLINE=97112961; PubMed=8943214;
RA Bieherich C.J., Fujita K., He W.-W., Jay G.;
RT "Prostate-specific and androgen-dependent expression of a novel
RT homeobox gene";
RL J. Biol. Chem. 271:31779-31782(1996).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster; TISSUE=Embryo;
RX MEDLINE=97287410; PubMed=9142502;
RA Sciavolino P.J., Abrams E.W., Yang L., Austenberg L.P., Shen M.M.,
RA Abate-Shen C.;
RT "Tissue-specific expression of murine Nkx3.1 in the male urogenital
RT system";
RL Dev. Dyn. 209:127-138(1997).
CC -1- FUNCTION: TRANSCRIPTION FACTOR, WHICH BINDS PREFERENTIALLY THE
CC CONSENSUS SEQUENCE 5'-TAAGTAC-3' AND CAN BEHAVE AS A
CC TRANSCRIPTIONAL REPRESSOR (BY SIMILARITY). PLAYS AN IMPORTANT ROLE

CC IN REGULATING PROLIFERATION OF GLANDULAR EPITHELIUM AND IN THE
CC FORMATION OF DUCTS IN PROSTATE AND MINOR SALIVARY GLANDS
CC (PARTICULARLY PALATINE AND LINGUAL GLANDS). ESSENTIAL FOR
CC APPROPRIATE DIFFERENTIATION AND SECRETORY FUNCTION OF THE
CC BUCCOPHARYNGEAL GLAND.
CC -1- SUBUNIT: INTERACTS WITH SERUM RESPONSE FACTOR (SRF).
CC -1- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED MOSTLY IN THE MALE UROGENITAL TRACT,
CC WITH HIGHEST EXPRESSION IN THE EPITHELIAL CELLS LINING THE DUCTS
CC OF ANTERIOR, DORSOLATERAL AND VENTRAL PROSTATE AND IN THE
CC BUCCOPHARYNGEAL GLAND, AND MUCH LOWER IN THE SEMINAL VESICLE AND THE
CC TESTIS. EXPRESSION IN THE PROSTATE INCREASES DURING SEXUAL
CC MATURATION AND IS DRASTICALLY REDUCED FOLLOWING CASTRATION.
CC EXPRESSED ALSO IN BRAIN (HIPPOCAMPUS AND EXTERNAL GRANULAR LAYER
CC OF THE CEREBRAL CORTEX), KIDNEY (INTRALOBULAR ARTERIES), THYMUS
CC AND ADRENAL AND SALIVARY GLANDS.
CC -1- DEVELOPMENTAL STAGE: EARLY MARKER OF THE SCLEROTOME AND OF A
CC SUBSET OF VASCULAR SMOOTH MUSCLE CELLS. EXPRESSED ALSO IN
CC OUTGROWTHS OF EPITHELIAL CELLS, IN ECTODERMAL EPITHELIAL CELLS AND
CC IN RESTRICTED REGIONS OF THE CENTRAL NERVOUS SYSTEM. DETECTED
CC FIRST AT 7.5 DPC IN THE PARAXIAL MESODERM ADJACENT TO THE NEURAL
CC FOLD. AT 8.5 DPC, SEGMENTAL EXPRESSION IN THE FIRST 8 OR 9
CC SOMITES. EXPRESSION PROCEEDS CAUDALLY IN PARALLEL WITH SOMITE
CC MATURATION AND IS RESTRICTED TO THE SCLEROTOME. AS THE SOMITES
CC MATURE, EXPRESSION MOVES AWAY FROM THE AXIAL STRUCTURES, BECOMES
CC TRANSIENTLY RESTRICTED TO A SUBSET OF EARLY MYOTOMAL CELLS AT THE
CC DORSAL MEDIAL LIP AND IS SUBSEQUENTLY DOWNREGULATED. AT 10.5 DPC,
CC EXPRESSED ONLY IN THE MOST CAUDAL IMMATURE SOMITES. AT 9.5 DPC,
CC PRESENT IN THE DORSAL AORTA. AT 11.5 DPC, RESTRICTED TO THE
CC VASCULAR SMOOTH MUSCLE CELLS OF CAUDAL REGION OF THE DORSAL AORTA.
CC AT 12.5 DPC, EXPRESSED IN THE DISTAL EPITHELIUM OF THE TONGUE AND
CC IN RAHKE'S POUCH (ANTERIOR PITUITARY). BY 13.5 DPC, ALSO DETECTED
CC IN TOOTH BUDS. EXPRESSION IN THE ABDOMINAL AORTA CONTINUES THROUGH
CC 11.5 TO 15.5 DPC. DETECTED IN THE VENTRAL VESSELS AT 12.5 DPC,
CC IN THE CAROTID VESSEL AT 13.5 DPC AND IN ARTERIES AND INTERLOBULAR
CC ARTERIES OF THE KIDNEY AT 15.5 DPC. IN NEONATES, PRESENT IN
CC PALATINE GLANDS, EPITHELIAL ROOT SHEATH OF THE TOOTH AND
CC EPITHELIAL HAIR SHEATH. IN THE NERVOUS SYSTEM OF NEONATES,
CC EXPRESSED IN THE OLFACTORY LOBE, OLFACTORY EPITHELIAL CELLS AND
CC CEREBELLAR CORTEX. EXPRESSED IN THE MALE UROGENITAL SYSTEM DURING
CC LATE EMBRYOGENESIS: AT DAY 14.5, EXPRESSED IN THE OUTBUDINGS OF
CC THE PELVIC REGION OF THE UROGENITAL SINUS, AND, AT LOWER LEVELS,
CC IN THE PROSPECTIVE URETHRA. EXPRESSION IS CONFINED TO THE
CC EPITHELIAL CELLS THAT ARE INVAGINATING INTO THE SURROUNDING
CC MESENCHYME, WITH HIGHEST LEVELS AT THE LEADING EDGE. AT 17.5 DPC,
CC PRESENT IN THE DEVELOPING VENTRAL, DORSOLATERAL AND ANTERIOR
CC PROSTATIC BUDS, IN THE NASCENT BUCCOPHARYNGEAL GLANDS, AS WELL AS IN
CC THE EPITHELIAL DUCTS THAT JOIN THE GLANDS TO THE PROSPECTIVE
CC URETHRA. DURING POSTNATAL GROWTH AND MORPHOGENESIS OF THE
CC PROSTATE, HIGH EXPRESSION IS MAINTAIN AT SITES OF DUCTAL OUTGROWTH
CC AND BRANCHING. IN THE DEVELOPING TESTIS, DETECTED AT 14.5 AND 17.5
CC DPC IN THE MEDULLARY CORDS, WHICH FORM SEMINIFEROUS TUBULES.
CC -1- INDUCTION: BY ANDROGENS. DURING EMBRYONIC DEVELOPMENT, INDUCED AND
CC MAINTAINED BY SONIC HEDGEHOG IN PRE-SOMATIC MESODERM, IN IMMATURE
CC SOMITES AND IN UROGENITAL SINUS, BUT NOT IN THE OTHER EXPRESSION
CC DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE NK-3 FAMILY OF HOMEBOX PROTEINS.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U73460; AAC52956.1; -
DR EMBL: U88542; AAB58025.1; -
DR HSSP: P23441; 1FTT.
DR GSD: MGI:97352; NKX3-1.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.

DR SMART: SM00389; HOX; 1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS0071; HOMEBOX_2; 1.
 KW Developmental protein; transcription regulation; repressor;
 KM DNA-binding; Nuclear protein; Homebox.
 FT DNA_BIND 125 184 HOMEBOX.
 SQ SEQUENCE 237 AA; 26824 MW; 4B074387F3BA1223 CRC64;

Query Match 11.4%; Score 70.5; DB 1; Length 237;
 Best Local Similarity 27.3%; Pred. No. 8.7;
 Matches 24; Conservative 10; Mismatches 35; Indels 19; Gaps 3;

QY 25 EPOEEVYVGGGR-----SKRDPDIVQLRLFRKSHSSLEGLKALSDPDKEST 75
 DB 6 EPEPEYVAGGRSPMAAPPQSKRLTSF--LIDILDRRAERHGHSGNPFQSDPRRDS 63
 QY 76 SPEKRDHDFVGLMGKRSYQDPSPTDV 103
 DB 64 APEP-----DKAGRGVAPEDPSI 83

RESULT 8
 MAT2.COCHIE STANDARD; PRT; 343 AA.
 ID MAT2.COCHIE
 AC 002991;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Mating-type protein MAT-2.
 GN MAT2.
 OS Cochliobolus heterostrophus (Drechslera maydis).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
 OC Pleosporales; Pleosporaceae; Cochliobolus.
 OX NCBI_TaxID=5016;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 48330 / C3;
 RX MEDLINE=93241164; PubMed=8479433;
 RA Turgeon B.G., Bohlmann H., Cluffetti L.M., Christiansen S.K.,
 RA Yang G., Schaefer W., Yoder O.C.,
 RT "Cloning and analysis of the mating type genes from Cochliobolus
 RT heterostrophus.";
 RL Mol. Genet. 238:270-284(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C4;
 RX MEDLINE=98420099; PubMed=9749670;
 RA Wiersel S., Horwitz B., Yamaguchi K., Yoder O.C., Turgeon B.G.,
 RT "Single mating type-specific genes and their 3' UTRs control mating
 RT and fertility in Cochliobolus heterostrophus.";
 RL Mol. Genet. 259:272-281(1998).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: CONTAINS 1 HMG BOX.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X68398; CAA48464.1; .
 CC EMBL: AF027687; AAB84004.1; .
 CC PIR: S34811; S34811.
 CC HSSP: P48436; 1SX9.
 DR InterPro: IPR000910; HMG_12_box.
 DR Pfam: PF00505; HMG_box; 1.
 DR SMART: SM00398; HMG; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein; Fertilization.
 FT DNA_BIND 131 199 HMG BOX.
 SQ SEQUENCE 343 AA; 38319 MW; 6DE7F4CGC2214103 CRC64;

Query Match 11.3%; Score 70; DB 1; Length 343;
 Best Local Similarity 31.0%; Pred. No. 15;
 Matches 22; Conservative 8; Mismatches 35; Indels 6; Gaps 1;

QY 6 LFTAILAFSLAOSFGAVCEPOEEVYVGGGRKRDPLYLQRLFRKSHSSLEGLKALS 65
 DB 54 LFGGLIVEHFRKRCVGECPVELTYMDGG-----DNYHTLVQPKNMRSPOVYSSPQ 107
 QY 66 QASTDPKESTS 76
 DB 108 SAOTSPSEQTS 118

RESULT 9
 E2F3.MOUSE STANDARD; PRT; 356 AA.
 ID E2F3.MOUSE
 AC 035261;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Transcription factor E2F3 (E2F-3) (Fragment).
 GN E2F3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SWISS ALBINO; TISSUE=Fibroblast;
 RX MEDLINE=98025478; PubMed=9376316;
 RA Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L.,
 RA Phillips R.A.;
 RT "Expression patterns of the E2F family of transcription factors
 RT during mouse nervous system development.";
 RL Mech. Dev. 66:13-25(1997).
 RN [2]
 RP DEVELOPMENTAL EXPRESSION.
 RX MEDLINE=97294022; PubMed=9149906;
 RA Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L.,
 RA Phillips R.A.;
 RT "Expression patterns of the E2F family of transcription factors
 RT during murine epithelial development.";
 RL Cell Growth Differ. 8:553-563(1997).
 CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY
 CC WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GCC,
 CC FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS
 CC ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE
 CC DR1FL/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE
 CC PROGRESSION FROM G1 TO S PHASE. E2F-3 BINDS SPECIFICALLY TO RB1
 CC PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER.
 CC -1- SUBUNIT: COMPONENT OF THE DR1FL/E2F TRANSCRIPTION FACTOR COMPLEX.
 CC BINDS COOPERATIVELY WITH DP-1 TO E2F SITES. INTERACTS WITH
 CC RETINOBLASTOMA PROTEIN RB1 AND RELATED PROTEINS (SUCH AS P107)
 CC THAT INHIBIT THE E2F TRANSACTIVATION DOMAIN.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC
 CC EXPRESSED IN BOTH VENTRAL AND DORSAL REGIONS OF THE SPINAL CORD
 CC FROM 13.5 DPC. ALSO EXPRESSED IN DORSAL ROOT AND CRANIAL GANGLIA
 CC IN 11.5-18.5 DPC EMBRYOS. ONLY LOW LEVELS OF EXPRESSION IN
 CC DEVELOPING BRAIN. IN THE DEVELOPING RETINA (15.5 DPC), EXPRESSION
 CC OF E2F-3 IS LOCALIZED TO THE GANGLION CELL LAYER. IN OTHER
 CC DEVELOPING TISSUES, EXPRESSED IN LIVER, LUNG AND HEART. WEAK
 CC EXPRESSION IN DEVELOPING KIDNEY AND SKELETAL MUSCLE. ABSENT FROM
 CC THE DEVELOPING CHOROID PLEXUS, THYMUS AND DEVELOPING SKIN. LOW
 CC MESENCHYME IN 12.5-18.5 DPC EMBRYOS.
 CC
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF015948; AAB71671.1; -
 DR HSSP; 016254; 1CF7.
 DR TRANSFAC; T02947; -
 DR MGD; MGI:1096340; E2F3
 DR InterPro; IPR003316; E2F_TDP.
 DR Pfam; PF02319; E2F_TDP; 1.
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein; Cell cycle.
 FT NON_TER 1 1
 FT DOMAIN <1 44 CYCLIN A/CDK2 BINDING (POTENTIAL).
 FT DNA_BIND 46 136 POTENTIAL.
 FT DOMAIN 95 116 LEUCINE-ZIPPER.
 FT DOMAIN 100 136 DEF BOX.
 FT DOMAIN 137 228 DIMERIZATION (POTENTIAL).
 FT DOMAIN 282 356 TRANSACTIVATION (POTENTIAL).
 FT DOMAIN 323 340 P130 PROTEIN ASSOCIATION (POTENTIAL).
 FT DOMAIN 323 340
 SO SEQUENCE 356 AA; 39007 MW; 07BFC4E0AC8E77 CRC64;

Query Match 11.1%; Score 69; DB 1; Length 356;
 Best Local Similarity 24.7%; Pred. No. 19;
 Matches 24; Conservative 17; Mismatches 42; Indels 14; Gaps 3;

QY 23 CKEP---QEEVVGSGRSKRDPDLYQ-----LQRLFKSHSLLEGKLLKLSQASTDPKE 73
 DB 1 CSSPRLKDPAPGRGGGGGPPAKRRLESGHQYLSGDKTPKGRGALRSPDSPRT 60
 QY 74 STSPKRMHPFVGLMGRSVQ-----PDSPDVNQ 105
 DB 61 PRSPKRTYDTSLGTLTKKFIQLLSQSPGVLDLNR 97

RESULT 10
 E2F3_HUMAN
 ID E2F3_HUMAN STANDARD; PRT; 465 AA.
 AC 000716; Q15000;
 DT 01-NOV-1997 (Rel. 35; Created)
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Transcription factor E2F3 (E2F-3).
 GN E2F3 OR KIAA0075.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94067142; PubMed=8246996;
 RA Lees J.A., Saito K., Vidal M., Valentine M., Look T., Harlow E.,
 RA Dyson N., Helin K.;
 RT "The retinoblastoma protein binds to a family of E2F transcription factors.";
 RL Mol. Cell. Biol. 13:7813-7825(1993).
 RN [2]
 RP SEQUENCE OF 292-465 FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=96051398; PubMed=7584044;
 RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
 RA Seki N., Kawatabayashi Y., Ishikawa K.-I., Tabata S.;
 RT "Prediction of the coding sequences of unidentified human genes. II. The coding sequences of 40 new genes (KIAA041-KIAA080) deduced by analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 1:223-229(1994).
 CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GCGC. FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE DRF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE

PROGRESSION FROM G1 TO S PHASE. E2F-3 BINDS SPECIFICALLY TO RB1 PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER.
 -1- SUBUNIT: COMPONENT OF THE DRF1/E2F TRANSCRIPTION FACTOR COMPLEX. BINDS COOPERATIVELY WITH DP-1 TO E2F SITES. INTERACTS WITH RETINOBLASTOMA PROTEIN RB1 AND RELATED PROTEINS (SUCH AS P107) THAT INHIBIT THE E2F TRANSACTIVATION DOMAIN.
 -1- SUBCELLULAR LOCATION: Nuclear.
 -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; Y10479; CAA71504.1; -
 DR EMBL; D38550; BAA07553.1; -
 DR HSSP; 016254; 1CF7.
 DR MIM; 600427; -
 DR InterPro; IPR003316; E2F_TDP.
 DR Pfam; PF02319; E2F_TDP; 1.
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein; Cell cycle.
 FT DOMAIN 26 31 POLY-ALA.
 FT DOMAIN 45 53 POLY-ALA.
 FT DOMAIN 101 153 CYCLIN A/CDK2 BINDING (POTENTIAL).
 FT DOMAIN 120 129 POLY-GLY.
 FT DOMAIN 155 245 POTENTIAL.
 FT DOMAIN 204 225 LEUCINE-ZIPPER.
 FT DOMAIN 209 245 DEF BOX.
 FT DOMAIN 246 337 DIMERIZATION (POTENTIAL).
 FT DOMAIN 391 465 TRANSACTIVATION (POTENTIAL).
 FT DOMAIN 432 449 RETINOBLASTOMA PROTEIN ASSOCIATION (POTENTIAL).
 SO SEQUENCE 465 AA; 49161 MW; 4641565842CA99EC CRC64;

Query Match 11.1%; Score 69; DB 1; Length 465;
 Best Local Similarity 28.0%; Pred. No. 26;
 Matches 23; Conservative 17; Mismatches 30; Indels 12; Gaps 3;

QY 33 GGGRSKRDPDL-----YQLLRFKSHSLLEGKLLKLSQASTPKSTSPKRDMDPFVG 88
 DB 128 GGPAPKRRLESGHQYISDGLKTP--KGKGRALRSPDSPKPSPKRTYDTSLG 184
 QY 89 LMGKRSVQ-----PDSPDVNQ 105
 DB 185 LTKKFIQLLSQSPGVLDLNR 206

RESULT 11
 BP28_DROME
 ID BP28_DROME STANDARD; PRT; 2096 AA.
 AC 09VM75;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Hypothetical protein CG10805.
 GN CG10805.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Gelinkler S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayan A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhanderi D., Boshakov S.,
 RA Boriva D., Botchan M.A., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., Mpherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Manzy D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Ruskern D.R., Paclad J.M.,
 RA Palazzolo M., Picman G.S., Pan S., Pollard J., Puri V., Shen M.G.,
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shee H.,
 RA Shue B.C., Slden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spletter E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtks R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Welstock G.M., Weltschbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: BELONGS TO THE BAR28 FAMILY.
 CC -1- SIMILARITY: CONTRAINS 1 HEAT REPEAT.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE003615; AAF52447.2;
 CC FLYBase: FBgn0031864; CG10805.
 CC InterPro: IPR001810; F-box.
 CC InterPro: IPR000357; HEAT_repeat.
 CC Pfam: PF00646; F-box; 1.
 CC PROSITE: PS50077; HEAT_REPEAT; FALSE_NEG.
 CC Hypothetical protein.
 CC REPEAT 2058 2094 HEAT.
 CC FT SEQUENCE 2096 AA; 237217 MW; 3E7B3C67CE6DF62C CRC64;
 SO

Query Match 11.1%; Score 69; DB 1; Length 2096;
 Best Local Similarity 35.2%; Pred. No. 1.4e+02;
 Matches 25; Conservative 14; Mismatches 18; Indels 14; Gaps 5;
 Oy 43 LYQLQRLFKSHSSLEGILKALQAST--DPKESTSPKRDMDHFFVGLMGRSVQPPSP 100
 Db 5 LAOQLQKLAAPSSVYTLADARSASILDFPKKATKDRSIVE--ISLTGQEL----- 56
 Oy 101 TDVNOENVPSF 111
 Db 57 TDEN---PAF 63
 RESULT 12
 TKNL_MESAU STANDARD; PRT: 130 AA.

AC 060541; P49110;
 DT 01-NOV-1997 (Rel. 35; Created)
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Proctachykinin 1 precursor (PPT) [Contains: Substance P; Neurokinin A
 DE (NKA) (Substance K) (Neurokinin L); Neuropeptide K (NPK); Neuropeptide
 DE gamma; C-terminal flanking peptide].
 GN TAC1 OR NKA OR TAC2 OR NKA.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_Taxid=10036;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS BETA AND GAMMA).
 RC STRAIN=AURA; TISSUE=Brain;
 RA Heitland A., Krinofer M., Juergen Meagert H.J., Forssmann W.G.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),
 CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X80662; CA556691.1;
 CC EMBL: X80663; CA556692.1;
 CC InterPro: IPR003580; Proctachykinin.
 CC InterPro: IPR002040; Tachykinin.
 CC Pfam: PF02202; Tachykinin; 1.
 CC ProDom: PD005598; Proctachykinin; 1.
 CC SMART: SM00203; TK; 2.
 CC PROSITE: PS00267; TACHYKININ; 2.
 CC KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
 CC Amdatation; Alternative splicing; Signal; Neurotransmitter.
 CC FT SIGNAL 1 19
 CC FT PROPEP 20 56
 CC FT PEPTIDE 58 68
 CC FT PEPTIDE 72 107
 CC FT PEPTIDE 72 107
 CC FT PEPTIDE 72 107
 CC FT PEPTIDE 89 107
 CC FT PEPTIDE 98 107
 CC FT PEPTIDE 111 126
 CC FT PEPTIDE 111 126
 CC FT MOD_RES 68 68
 CC FT MOD_RES 107 107
 CC FT MOD_RES 107 107
 CC FT VARSPLIC 74 88
 CC FT SEQUENCE 130 AA; 14907 MW; CC92E9371A646F2E CRC64;
 SO

Query Match 11.0%; Score 68; DB 1; Length 130;
 Best Local Similarity 27.0%; Pred. No. 7.8;
 Matches 33; Conservative 13; Mismatches 36; Indels 40; Gaps 6;
 Oy 4 MLFPAIAPFSL-----AQSGA-----YKPEEEVPPGGGRKR 39
 Db 1 MKILYAVAFVFLVYQLSLEIGANDLNYMSDSQIKALPEPEPHILQRIARPK 60
 Oy 40 DDDLYQLQRLFKSHSSLE---GLLKAL---SQASTDPKESTSPKRDMDHFFVGLMGR 93
 Db 61 POFEGGLMGRK-DADNSTEKOVALLKALYGHGQIS-----HKRKTSDFVGLMGR 110
 Oy 94 SV 95
 Db 111 AL 112

```

RESULT 13
RDGC_XYLFA STANDARD; PRT; 302 AA.
ID RDGC_XYLFA
AC 09PT9;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Recombination associated protein rdgc.
GN RDGC OR XF0568.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=3371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Stimpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Klieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos J.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marquês M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peloto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
RA Quagaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E.Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A.Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Valada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Secubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
CC -1- FUNCTION: NOT KNOWN; MAY BE INVOLVED IN RECOMBINATION (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Seems to be associated with the nucleoid (by
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE RDGC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF003903; AAF83378.1; -
CC DNA recombination; Complete proteome.
CC KW SEQUENCE 302 AA; 34070 MW; 705322F88F681B3 CRC64;

```

Query Match 11.0%; Score 68; DB 1; Length 302;

Best local Similarity 31.2%; Pred. No. 20;

Matches 24; Conservative 11; Mismatches 32; Indels 10; Gaps 3;

```

QY 23 CKEPQEEV--VPGGGRSKRPD--LYQLRLFKSHSLBGLK-----ALSOASTDPK 72
I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
DB 86 CSIEEKGRHPGGRERRRIKDLIHLLPRAFAVKKSRIDAMDLRYGYAVDVASAKAA 145

```

```

QY 73 ESTSPKRMHDFEYGL 89
I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
DB 146 ETVISEIDLIGSPFAL 162

RESULT 14
CXAS_RAT STANDARD; PRT; 355 AA.
ID CXAS_RAT
AC P28234;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gap junction alpha-5 protein (Connexin 40) (Cx40).
GN GJA5 OR CXN-40.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=92112940; PubMed=1370487;
RA Haefliger J.-A., Bruzone R., Jenkins N.A., Gilbert D.J.,
RA Copeland N.G., Paul D.L.;
RT "Four novel members of the connexin family of gap junction proteins.
RT Molecular cloning, expression, and chromosome mapping."
RL J. Biol. Chem. 267:2057-2064(1992).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021055; PubMed=1328644;
RA Beyer E.C., Reed K.E., Westphale E.M., Kanter H.L., Larson D.M.;
RT "Molecular cloning and expression of rat connexin40, a gap junction
RT protein expressed in vascular smooth muscle."
RL J. Membr. Biol. 127:69-76(1992).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98066770; PubMed=9403066;
RA Seul K.H., Tadros P.N., Beyer E.C.;
RT "Mouse connexin40: gene structure and promoter analysis."
RL Genomics 46:120-126(1997).
CC -1- FUNCTION: ONE GAP JUNCTION CONSISTS OF A CLUSTER OF CLOSELY PACKED
CC PAIRS OF TRANSMEMBRANE CHANNELS, THE CONNEXONS, THROUGH WHICH
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBORING CELL.
CC -1- SUBUNIT: A CONNEXON IS COMPOSED OF A HEXAMER OF CONNEXINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG.
CC -1- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. ALPHA-TYPE (GROUP II)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M76535; AAA41000.1; -
CC EMBL: M83092; AAA41194.1; -
CC DR EMBL: AF021806; AAC53502.1; -
CC DR EMBL: AF022136; AAC53503.1; -
CC DR PIR: A42053; A42053.
CC DR InterPro: IPR000500; Connexin.
CC DR Pfam: PF00029; connexin; 1.
CC DR PRINTS: PR00206; CONNEXIN.
CC DR SMART: SM00037; CNX; 1.
CC DR PROSITE: PS00407; CONNEXINS_1; 1.
CC DR PROSITE: PS00408; CONNEXINS_2; 1.
CC KW Gap junction; Transmembrane.
CC FT INT MET 0 18 BY SIMILARITY.
CC FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 19 39 POTENTIAL.

```

FT DOMAIN 40 75 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 76 96 POTENTIAL.
FT DOMAIN 97 162 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 163 183 POTENTIAL.
FT DOMAIN 184 203 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 204 224 POTENTIAL.
FT DOMAIN 225 355 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 355 AA; 40102 MW; 4F85943868600B9 CRC64;

Query Match 10.9%; Score 67.5; DB 1; Length 355;
Best Local Similarity 28.4%; Pred. No. 27;
Matches 25; Conservative 14; Mismatches 34; Indels 15; Gaps 4;

OY 42 DLYOL-----LORLEKSHS-----SLEGLKALSOASTDPKE-----STSPKRDHDFEV 87
DB 223 ELIHLGKWKIKRRLAKSKQGDKHOLGPTSLVQGLTPPPDNOCLKNSPDKFSDSN 282
OY 88 GLMGRSVQPDSPDTDV-NOENVSPFGIL 114
DB 283 NMGSRKNPDPLATEVPPNQEIPEEGFI 310

RESULT 15
VCAP_HSV11 STANDARD; PRT; 1374 AA.
ID VCAP_HSV11
AC P06491;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major capsid protein (MCP) (Capsid protein VP5).
GN UL19.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
CX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
herpes simplex virus type 1."
RT J. Gen. Virol. 69:1531-1574(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87010565; PubMed=3020164;
RA Davison B.A.J., Scott J.E.;
RT "DNA sequence of the major capsid protein gene of herpes simplex
virus type 1."
RT J. Gen. Virol. 67:2279-2286(1986).
RN [3]
RP SEQUENCE OF 202-211 AND 607-616.
RX MEDLINE=93019027; PubMed=1328483;
RA Davison M.D., Rixon F.J., Davison A.J.;
RT "Identification of genes encoding two capsid proteins (VP24 and VP26)
of herpes simplex virus type 1."
RT J. Gen. Virol. 73:2709-2713(1992).
RL J. Gen. Virol. 73:2709-2713(1992).
CC -1- SUBUNIT: MAJOR PROTEIN OF THE ICOSAEDRAL CAPSID.
CC -1- FUNCTION: MAIN COMPONENT OF THE HEXAVALENT, AND PROBABLY THE
PERTAVALENT CAPSOMERES.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES MAJOR CAPSID PROTEIN
FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D10879; BAA01665.1; -.

DR EMBL; X14112; CAA32332.1; -.
DR EMBL; X04467; CAA28154.1; -.
DR PIR; A27239; VCB817.
DR PIR; A30084; A30084.
DR InterPro; IPR000912; Herpes_MCP.
DR Pfam; PF03122; Herpes_MCP; 1.
DR PRINTS; PR00235; HSVCAPSIDMCP.
KW Coat protein.
SQ SEQUENCE 1374 AA; 149083 MW; 859C76E2EADDE05B7 CRC64;

Query Match 10.9%; Score 67.5; DB 1; Length 1374;
Best Local Similarity 21.7%; Pred. No. 1.2e+02;
Matches 28; Conservative 17; Mismatches 43; Indels 41; Gaps 5;

OY 10 ILAFSLASFGAVCKEPEEYVPGGGRKRDP-----DLYOLL----- 47
DB 449 VLRLSLERHAIGVCHPSLMDAAVGLNRDVEANPYGAVAAPAGPAAQMQLFLNA 508
OY 48 --QRLFKSHSLEGLKALSOASTDPKESTSPKRDH-----HDFEVL-----MGKRS 94
DB 509 MGQRL-----AHGRVHWVAEGQMPPEQFMQPDNANLALHLPARDFVGVADVELPGSD 562
OY 95 VOPDSPDTDV 103
DB 563 VPPAGPGEI 571

Search completed: May 24, 2002, 17:05:34
Job time: 118 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 17:03:01 ; Search time 16.34 Seconds

(without alignments)
711.555 Million cell updates/sec

Title: US-09-852-659-68

Perfect score: 619

Sequence: 1 MRIMLFTAILAFSLAQSFSG.....DVNQENVPSFGILKYPRAE 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	416.5	67.3	126	2	A25905
2	373.5	60.3	116	2	tachykinin B precu
3	353.5	57.1	116	2	neurokinin B precu
4	77.5	12.5	2380	2	T29551
5	77	12.1	2211	1	KRE05
6	75	12.1	1008	2	T05578
7	72	11.6	921	2	S35160
8	72	11.6	969	2	A75634
9	70.5	11.4	1711	2	T21432
10	70	11.3	343	2	S34811
11	69.5	11.2	271	2	T40982
12	69	11.1	343	2	S66173
13	69	11.1	742	2	D86160
14	68.5	11.1	437	2	T26767
15	68.5	11.1	807	2	T19836
16	68.5	11.1	839	2	T45908
17	68	11.0	130	2	S47038
18	68	11.0	202	2	AD2197
19	68	11.0	252	2	AC2226
20	68	11.0	302	2	G82791
21	68	11.0	808	2	G86185
22	68	11.0	906	2	T47340
23	68	11.0	3076	2	A87058
24	67.5	10.9	356	1	A42053
25	67.5	10.9	1374	1	VCBE17
26	67.5	10.9	2464	1	QRMSP1
27	67	10.8	932	2	T21338
28	67	10.8	1028	2	A56038
29	67	10.8	1213	2	S16356

30	66.5	10.7	193	2	S70379	outer surface prot
31	66.5	10.7	773	2	G83816	late competence op
32	66.5	10.7	978	1	RGBY13	regulatory protein
33	66.5	10.7	1126	2	JC4019	DNA mismatch repa
34	66	10.7	130	1	SPRB	substance P beta p
35	66	10.7	293	1	B4502	nonstructural prot
36	66	10.7	452	2	S72266	translation initia
37	66	10.7	823	2	T16758	hypothetical prote
38	66	10.7	1734	2	A54602	microtubule-associ
39	65.5	10.6	129	1	SPHUB	neurokinin 1 precu
40	65.5	10.6	130	1	SPBOB	neurokinin 1 precu
41	65.5	10.6	450	1	VHVRV	nucleoprotein - ra
42	65.5	10.6	578	2	T48795	origin recognition
43	65.5	10.6	803	2	B82045	aspartokinase II/h
44	65.5	10.6	910	1	S73361	dnaf homolog prote
45	65.5	10.6	920	2	S53961	hypothetical prote

ALIGNMENTS

RESULT 1
A25905
tachykinin B precursor - bovine
N:Alternate names: neuromedin K
C:Species: Bos primigenius taurus (cattle)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A25905
R:Kotani, H.; Hoshimaru, M.; Nawa, H.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 83, 7074-7078, 1986
A>Title: Structure and gene organization of bovine neuromedin K precursor.
A:Reference number: A25905; M01D:86313713
A:Accession: A25905
A:Molecule type: DNA
A:Residues: 1-126 <KOT>
A:Cross-references: GB:M14351; NID:g163587; PIDN:AAA30723.1; PID:g163590
C:Superfamily: neurokinin B precursor
F:1-26/Domain: signal sequence #status:predicted <SIG>
F:27-126/Product: protachykinin B #status:predicted <MAT>

Query Match
Best Local Similarity 67.3%; Score 416.5; DB 2; Length 126;
Matches 83; Conservative 15; Mismatches 23; Indels 5; Gaps 2;
QY 1 MRIMLFTAILAFSLAQSFQAVCKEPEQEVDPGGRKRDPLYO----LQRLFKSH-S 55
DB 1 MSSTLFFAVIALSSARSLGAVCEESQEQVYPGGGSHKSDMLXOLPSSLRLRLYDSRVY 60
QY 56 SLEGLKALSOASTDPKKESTPEKRDMDHDFVGLMGKRSVOPDPTDVNQENVPSFGILK 115
DB 61 SLDGLKMLSKASVSGKESPDPKRDMDHDFVGLMGKRLMDPTVDVNDQENVPSFGITK 120
QY 116 YPPRAE 121
DB 121 YPPSVE 126
RESULT 2
I65342
tachykinin B precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I65342
R:Take, K.; Munkata, E.; Hosaka, M.; Murakami, K.; Nakayama, K.
Biomed. Res. 14, 253-259, 1993
A>Title: Cloning and sequence analysis of mouse cDNAs encoding preprotachykinin A and
A:Reference number: I52526
A:Accession: I65342
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-116 <RES>
A:Cross-references: GB:D14423; NID:g407346; PIDN:BA003316.1; PID:g407347

Best Local Similarity 28.6%; Pred. No. 52;
Matches 32; Conservative 13; Mismatches 39; Indels 28; Gaps 6;

OY 14 SLAOSFVACPEP---QEEVPPGGGRKRDPLDY-----LLORLFKSHSLGLKAL 64

Db 1269 SLSPDLGOTATSPPGQSLSPDLGOTATSPDLQESLSPDLGOTATSPDPSQSLSPDL 1328

OY 65 SOASNDP---KESTSPKRDHDFVGLMKRSYVDPSPVNOENV-PSFG 112

Db 1329 GOTALSPDPSQSLSPD-----LGOTSLSP---DLGQESLSPDLG 1365

RESULT 6

hypothenical protein F22K18.240 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999

C:Accession: T05578

R:Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hohnsels, J.; Mewes, H.W.; Mayer, K.F.

A:Reference number: 215419

A:Accession: T05578

A:Molecule type: DNA

A:Residues: 1-1008 <BEV>

A:Cross-references: EMBL:AL035356

A:Experimental source: cultivar Columbia; BAC clone F22K18

C:Genetics:

A:Map position: 4

A:Introns: 94/2; 521/3; 550/2; 579/1; 633/2; 674/3; 727/2; 755/3; 825/3; 846/2

A:Note: F22K18.240

Query Match

Best Local Similarity 12.1%; Score 75; DB 2; Length 1008;

Matches 29; Conservative 8; Mismatches 41; Indels 10; Gaps 4;

OY 32 PGGGSKRDPDIYOLQRLFK-----SHSLEGLKALSQASTDPKESPEKRMHDF 86

Db 896 PGGGSGNIGSFYSFORLQKLEEDSASDSSSLFDSNSDECSSTDSMD--DFADFI 953

OY 87 VG-LMGKRSVDPD--SPTDVNOENVPSF 111

Db 954 FGDHGRAGHGSPTSPSSSSSPPF 981

RESULT 7

aspartate kinase (EC 2.7.2.4) / homoserine dehydrogenase (EC 1.1.1.3) precursor - carrot

C:Species: Daucus carota (carrot)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999

C:Accession: S35160; S38836

R:Weisemann, J.M.; Matthews, B.F.

A:Reference number: S35160; MUID:93283634

A:Accession: S35160

A:Molecule type: mRNA

A:Residues: 1-921 <WEI>

A:Cross-references: EMBL:L11529; NID:g167547; PIDN:AAAI6972.1; PID:g464225

A:Accession: S38836

A:Molecule type: Protein

A:Residues: 338-358;405-433,'S',435-437;698-726,739,'A',740-741,'LGSPS' <WE2>

C:Superfamily: thra bifunctional enzyme; aspartate kinase homology; homoserine dehydroge

C:Keywords: chloroplast; multifunctional enzyme; oxidoreductase; phosphotransferase

F:92-557/Domain: aspartate kinase homology <DK1>

F:560-818/Domain: homoserine dehydrogenase homology <HSD>

F:564-592/Region: beta alpha beta NAD(P) nucleotide-binding fold

Query Match 11.6%; Score 72; DB 2; Length 921;

Best Local Similarity 31.2%; Pred. No. 59;

Matches 29; Conservative 13; Mismatches 41; Indels 10; Gaps 3;

OY 30 VPPGGRKRDPLDYOLQRLFKSHSLF-GLKALSQASTDPKESPEKRMHDFYNG 88

Db 243 VVNPAGSNQVDPDLFSEKRLKFWSSNOCOTIVATGFIASTPONTITLKRDSDFSA 302

OY 89 LMGR-KSVQDPSPTDVNOENVPSFGLKTPPR 119

Db 303 IMGALLRAGQVITWTDVN-----GVYSADPR 328

RESULT 8

McB-related protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: A75634

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.

A:Reference number: A75250; MUID:20036896

A:Accession: A75634

A:Molecule type: DNA

A:Residues: 1-969 <WHI>

A:Cross-references: GB:AE01826; NID:96460827; PIDN:AAF12568.1; PID:96460864; TIGR:DR

A:Experimental source: strain R1

C:Genetics:

A:Map position: megaplasmid

A:Genome: plasmid

A:Note: plasmid MPI

Query Match

Best Local Similarity 11.6%; Score 72; DB 2; Length 969;

Matches 33; Conservative 14; Mismatches 54; Indels 24; Gaps 4;

OY 7 FTALAFSLAOSFGAVCKEPEEVVPGGGRK-----DPDLY-QLQRLFKS 53

Db 368 FGALFLPLGGRFAEAQOQYMSGVRRLRVGLSLPERHLESPPELDAALTY 427

OY 54 HSLSEGLKALSQASTDPKESPEKRMHDFVGLMKRSVOPDSPDVNOENVPSFGI 113

Db 428 LDALVDLDGLMR-----PSEETSTEE-----IGVETRVAPLPAAFPVPGVPLNQI 476

OY 114 LKYP 118

Db 477 LKGP 481

RESULT 9

hypothenical protein F26H11.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21432

R:Barlow, K.

A:Reference number: Z19421

A:Accession: T21432

A:Molecule type: translated from GB/EMBL/DBJ

A:Residues: 1-1711 <WLL>

A:Cross-references: EMBL:Z81515; PIDN:CA804197.1; GSPDB:GN00020; CESP:F26H11.2

A:Experimental source: clone F26H11

C:Genetics:

A:Gene: CESP:F26H11.2

A:Map position: 2

A:Introns: 56/3; 154/2; 373/2; 417/3; 505/2; 596/1; 703/2; 847/3; 922/3; 1034/2; 1078

Query Match 11.4%; Score 70.5; DB 2; Length 1711;

```

Best Local Similarity 25.6%; Pred. No. 1.7e+02;
Matches 30; Conservative 16; Mismatches 44; Indels 27; Gaps 5;

QY 4 MLFTALIAS---LAOSFGAVCKEPOEEVPGGGRSKRDPDLYQLLQRLFKSHSSLEG 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 28 LMFYLLMLSDCHFSQILFLQRPYDMAPRGRSKR-----KHPSSG- 73

QY 61 LKALSOASTDPKESTPE-----KRDMD-FFVGLMGKRSVOPDSPTDVQENV 108
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 -NSSINDEPSESTSSRRPRSRRLPKRYFDGDISPPPTKKRAQRETPSDAEVEV 129

RESULT 10
S34811
mating factor MAT1-2 - fungus (Cochliobolus heterostrophus),
C:Species: Cochliobolus heterostrophus, Bipolaris maydis
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 24-Sep-1999
C:Accession: S34811; S28071
R:Turgeon, B.G.; Bohlmann, H.; Cluffett, L.M.; Christiansen, S.K.; Yang, G.; Schaefer,
Mol. Gen. Genet. 238, 270-284, 1993
A:Title: Cloning and analysis of the mating type genes from Cochliobolus heterostrophus.
A:Reference number: S34810; MUID:93241164
A:Accession: S34811
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-343 <TUR>
A:Cross-references: EMBL:X68398; NID:g2585; PIDN:CAA48464.1; PID:g2586
C:Genetics:
A:Gene: MAT1-2
A:Introns: 162/1
C:Superfamily: unassigned HMG box proteins; HMG box homology
C:Keywords: DNA binding; nucleus; transcription regulation
F:129-203/Domain: HMG box homology <HMG1>

Query Match 11.3%; Score 70; DB 2; Length 343;
Best Local Similarity 31.0%; Pred. No. 29;
Matches 22; Conservative 8; Mismatches 35; Indels 6; Gaps 1;

QY 6 LFTALIASFGAVCKEPOEEVPGGGRSKRDPDLYQLLQRLFKSHSSLEG 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 54 LFGILLVHFHRCVGEFPELVYMDG-----DNHTLVQMPKNNMRSPQVYSSQ 107

QY 66 QASTDPKESTS 76
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 SAOTSPSEQTS 118

RESULT 11
T40982
hypothetical protein SPCC285.02c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
C:Accession: T40982; T41248
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Voiclaert, G.
submitted to the EMBL Data Library, October 1998
A:Reference number: 221961
A:Accession: T40982
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-271 <LYN>
A:Cross-references: EMBL:AL031966; PIDN:CAA21451.1; GSPDB:GN00068; SPDB:SPCC1442.17c
R:Seeger, K.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A:Reference number: 221981
A:Accession: T41248
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-271 <SE>
A:Cross-references: EMBL:AL031945; PIDN:CAA20841.1; GSPDB:GN00068; SPDB:SPCC285.02c
A:Experimental source: strain 972h-; cosmid c285
C:Genetics:

```

```

A:Gene: SPCC1442.17c; SPCC285.02c
A:Map position: 3
A:Introns: 111/3; 138/3

Query Match 11.2%; Score 69.5; DB 2; Length 271;
Best Local Similarity 23.2%; Pred. No. 25;
Matches 38; Conservative 17; Mismatches 55; Indels 41; Gaps 6;

QY 1 MRIMLFTALIASFGAVCKEPOEEV-----PGGRSKRP-----DLXQ 45
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 LQIQKLASRIEILRQKEALAKARRNVALGLKSPALAKRIEPLIMODIYELLE 63

QY 46 LQO-----RLFKSHSSLEGKALSOASTDPKESTSPKRDMDHFFVGLM 90
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 LQVVEILANRCVYLEKRAFNDMSFSSLYHVAAA---POLQIKELRFVHDFLVKLY 120

QY 91 GR---RSVQPSPTDVOENVPSFGLKYP 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 GKEFARLSPDLAT-----NOTAFYQLLYPP 146

RESULT 12
S66173
mating factor MAT-2 homolog - Bipolaris sacchari
N:Alternate names: DNA-binding protein MAT-2
C:Species: Bipolaris sacchari
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Sep-1999
C:Accession: S66173
R:Sharon, A.; Yamaguchi, K.; Christiansen, S.; Horvitz, B.A.; Yoder, O.C.; Turgeon, B
Mol. Gen. Genet. 251, 60-68, 1996
A:Title: An asexual fungus has the potential for sexual development.
A:Reference number: S66173; MUID:96213991
A:Accession: S66173
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-343 <SH>
A:Cross-references: EMBL:X95814; NID:g1212792; PIDN:CAA65081.1; PID:e223605; PID:g121
A:Experimental source: strain 764-1
C:Genetics:
A:Gene: MAT-2
A:Introns: 162/1
C:Superfamily: unassigned HMG box proteins; HMG box homology
C:Keywords: DNA binding
F:128-203/Domain: HMG box homology <HMG1>

Query Match 11.1%; Score 69; DB 2; Length 343;
Best Local Similarity 28.8%; Pred. No. 37;
Matches 23; Conservative 11; Mismatches 28; Indels 18; Gaps 2;

QY 9 ALLFSLASFGAVCKE-----POEEVPGGGRSKRDPDLYQLLQRLFKSHSS 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 ALLDNIQLFGLVHFHRCVGEFPELVYMDG-----DNHTLVQMPKNNTR 98

QY 57 LEGLKALSOASTDPKESTS 76
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 99 SPQVYSSLSQASTSPSEQTS 118

RESULT 13
D86160
hypothetical protein F22D16.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: D86160
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Matzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 17:00:21 ; Search time 12.99 Seconds
(without alignments)
227.521 Million cell updates/sec

Title: US-09-852-659-68

Perfect score: 1 MRIMLFTALIAFSIAQSG.....DVENQNPSPGILKYPRAE 121

Sequence: 1 MRIMLFTALIAFSIAQSG.....DVENQNPSPGILKYPRAE 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	601.5	97.2	122	2	US-08-879-995A-1
2	601.5	97.2	122	3	US-09-215-096-1
3	416.5	67.3	126	2	US-08-879-995A-3
4	416.5	67.3	126	3	US-09-215-096-3
5	353.5	57.1	116	2	US-08-879-995A-4
6	353.5	57.1	116	3	US-09-215-096-4
7	72	11.6	863	2	US-08-380-182-19
8	72	11.6	863	3	US-08-380-182-20
9	69	11.1	130	6	5268359-2
10	67.5	10.9	382	3	US-08-362-740-70
11	67.5	10.9	382	4	US-09-109-879-70
12	66.5	10.7	2353	4	US-08-984-709A-50
13	65.5	10.6	126	6	5268359-5
14	64.5	10.4	647	2	US-08-770-761A-8
15	64.5	10.4	705	2	US-08-770-761A-7
16	64.5	10.4	3119	2	US-08-457-273B-8
17	64.5	10.4	3119	1	US-08-246-982A-16
18	64.5	10.4	3119	1	US-08-453-265-16
19	64	10.3	284	1	US-08-390-858B-36
20	63.5	10.3	1005	2	US-08-935-450-2
21	63.5	10.3	3144	1	US-08-246-982A-6
22	63.5	10.3	3144	1	US-08-453-265-6
23	63.5	10.3	3144	2	US-08-457-273B-42
24	63.5	10.3	3144	3	US-08-556-419-21
25	63.5	10.3	3144	4	US-09-041-886-15
26	63	10.2	348	4	US-08-953-041-2
27	63	10.2	348	4	US-09-159-417-2

28	62	10.0	444	3	US-09-221-235-14	Sequence 14, Appl
29	62	10.0	444	3	US-09-221-928-14	Sequence 14, Appl
30	62	10.0	444	4	US-09-221-527-14	Sequence 14, Appl
31	62	10.0	444	4	US-09-221-236-14	Sequence 14, Appl
32	62	10.0	444	4	US-09-221-416-14	Sequence 14, Appl
33	62	10.0	444	4	US-09-221-245-14	Sequence 14, Appl
34	62	10.0	444	4	US-09-163-115-14	Sequence 14, Appl
35	62	10.0	444	4	US-09-221-528-14	Sequence 14, Appl
36	62	10.0	444	4	US-09-593-553-14	Sequence 14, Appl
37	62	10.0	444	4	US-09-221-237-14	Sequence 14, Appl
38	62	10.0	1360	4	US-09-393-569-2	Sequence 5, Appl1
39	62	10.0	2517	2	US-08-801-263A-5	Sequence 5, Appl1
40	62	10.0	2517	3	US-08-102-248-5	Sequence 6, Appl
41	61.5	9.9	406	3	US-08-362-740-68	Sequence 6, Appl
42	61.5	9.9	406	4	US-09-109-879-68	Sequence 6, Appl
43	61.5	9.9	618	2	US-08-770-761A-3	Sequence 2, Appl1
44	61.5	9.9	618	2	US-08-770-761A-2	Sequence 2, Appl1
45	61.5	9.9	662	2	US-08-770-761A-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-08-879-995A-1
Sequence 1, Application US/08879995A
Patent No. 5985606
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HUMAN PREPROTRYPKININ B
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,995A
FILING DATE: Herewith
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0326 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 122 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT03
CLONE: 2109906
US-08-879-995A-1
Query Match 97.2%; Score 601.5; DB 2; Length 122;

Best Local Similarity 97.5%; Pred. No. 2,9e-66; Indels 1; Gaps 1;
Matches 119; Conservative 2; Mismatches 0;

QY 1 MRIMLFTAILAFSLAOSFGAVCKEPOEEVYPGGGRSKRDPDLYQLQRLFKSHSLEGL 60
DB 1 MRIMLFTAILAFSLAOSFGAVCKEPOEEVYPGGGRSKRDPDLYQLQRLFKSHSLEGL 60

QY 61 LKALSOASTDPKESTSPEKRDMDHDFVGLMGKRSVOPDSPTDV-NOENVPSEGLTKYPPR 119
DB 61 LKALSOASTDPKESTSPEKRDMDHDFVGLMGKRSVOPDSPTEKXNOENVPSEGLTKYPPR 120

QY 120 AE 121
DB 121 AE 122

RESULT 2
US-09-215-096-1
Sequence 1, Application US/09215096
Patent No. 6008194
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/215,096
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/879,995
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0326 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 122 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUR03
CLONE: 2109906
US-09-215-096-1

Query Match 97.2%; Score 601.5; DB 3; Length 122;
Best Local Similarity 97.5%; Pred. No. 2,9e-66;
Matches 119; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 MRIMLFTAILAFSLAOSFGAVCKEPOEEVYPGGGRSKRDPDLYQLQRLFKSHSLEGL 60
DB 1 MRIMLFTAILAFSLAOSFGAVCKEPOEEVYPGGGRSKRDPDLYQLQRLFKSHSLEGL 60

QY 61 LKALSOASTDPKESTSPEKRDMDHDFVGLMGKRSVOPDSPTDV-NOENVPSEGLTKYPPR 119
DB 61 LKALSOASTDPKESTSPEKRDMDHDFVGLMGKRSVOPDSPTEKXNOENVPSEGLTKYPPR 120

QY 120 AE 121
DB 121 AE 122

RESULT 3
US-08-879-995A-3
Sequence 3, Application US/08879995A
Patent No. 5985606
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,995A
FILING DATE: Herewich
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0326 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 163590
US-08-879-995A-3

Query Match 67.3%; Score 416.5; DB 2; Length 126;
Best Local Similarity 65.9%; Pred. No. 1,5e-43;
Matches 83; Conservative 15; Mismatches 23; Indels 5; Gaps 2;

QY 1 MRIMLFTAILAFSLAOSFGAVCKEPOEEVYPGGGRSKRDPDLYQLQRLFKSH-S 55
DB 1 MRIMLFTAILAFSLAOSFGAVCKEPOEEVYPGGGRSKRDPDLYQLQRLFKSHSRYV 60

QY 56 SLEGLKALSOASTDPKESTSPEKRDMDHDFVGLMGKRSVOPDSPTDVNOENVPSEGLTK 115
DB 61 SLGLKMLSKASVGRKESPLPKRDMHDFVGLMGKRNLPDTPVDINOENVPSEGLTK 120

QY 116 YPPRAE 121
DB 121 YPPSVE 126

```

APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Kaser, Matthew R
TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,995A
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0326 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 205725
US-08-879-995A-4

Query Match 57.1%; Score 353.5; DB 2; Length 116;
Best Local Similarity 61.7%; Pred. No. 6.9e-36;
Matches 74; Conservative 16; Mismatches 21; Indels 9; Gaps 3.

QY 1 MRIMLFTAILAFLSAQSGAVCKEPOEEVYPGGSKRDPDLY---QLQRLFFSHS- 55
  ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 116
Db 1 MRSAMLEFAVLAISLMTFGACCEEPQEQ---GGRLSKDSDSLPLPPLRLTDSRSI 56

QY 56 SLEGLIKLAVSKASVGRKETSPLQKRMHDFVGLMGRKRSVQPSPTDVQENVPSSGIIL 115
  |||||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|||| 116
Db 57 SLEGLIKLAVSKASVGRKETSPLQKRMHDFVGLMGRKRSVQPSPTADVVEENTPSSGIIL 116

RESULT 6
US-09-215-096-4
; Sequence 4, Application US/09215096
; Patent No. 6006194
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA

```

RESULT 8
 US-08-380-182-20
 : Sequence 20, Application US/08380182
 : Patent No. 5858749
 :
 : GENERAL INFORMATION:
 : APPLICANT: Mathews, Benjamin F.
 : APPLICANT: Weismann, Jane M.
 : TITLE OF INVENTION: A Bifunctional Protein From Carrots
 : TITLE OF INVENTION: (Daucus carota) with Aspartokinase and Homoserine
 : TITLE OF INVENTION: dehydrogenase Activities"
 : NUMBER OF SEQUENCES: 23
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Janelle S. Graeter
 : STREET: Room 411, Bldg. 005, BARC-W
 : CITY: Beltsville
 : STATE: Maryland
 : COUNTRY: USA
 : ZIP: 20705
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/380,182
 : FILING DATE:
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Graeter, Janelle S.
 : REGISTRATION NUMBER: 35,024
 : REFERENCE/DOCKET NUMBER: 0226_94
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 301-504-6629
 : TELEFAX: 301-504-5060
 : INFORMATION FOR SEQ ID NO: 20:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 863 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : HYPOTHEICAL: NO
 : FRAGMENT TYPE: N-terminal
 : ORIGINAL SOURCE:
 : ORGANISM: Daucus carota

SEQUENCE CHARACTERISTICS:
 LENGTH: 382 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-109-879-70

Query Match 10.9%; Score 67.5; DB 4; Length 382;
 Best Local Similarity 25.4%; Pred. No. 4.4;
 Matches 33; Conservative 15; Mismatches 27; Indels 55; Gaps 7;

QY 11 LAFSL-----AOSFGAVCKEPEEVVPGGSRKRPD-LYOLLQR----- 49
 DB 147 LSFSLVGGQLEGAQALGPVEQEEEE--PG-----DEVLFRTETWRRLPPLCTPPV 199
 QY 50 -----LFKSHSLGLKALSOASTDPKESTSPKRDHDFVGLMKRSVQ 96
 DB 200 PALYCATRRLPGLLSHQAIPLH--SPTSPEPDPTSPESPDP-----TTS 245
 QY 97 PDSPTDVNOE 106
 DB 246 PDSPTTSOE 255

RESULT 12
 US-08-984-709A-50

; Sequence 50, Application US/08984709A
 ; Patent No. 6320032

; GENERAL INFORMATION:
 ; APPLICANT: Williams, Mark E.
 ; APPLICANT: Stauderman, Kenneth A.
 ; APPLICANT: Harpold, Michael M.
 ; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
 ; TITLE OF INVENTION: METHODS
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Heller Ehrman White & McCauliffe
 ; STREET: 4250 Executive Square, Suite 700
 ; CITY: La Jolla
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 92037

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/984,709A
 ; FILING DATE: 02-DEC-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie L.
 ; REGISTRATION NUMBER: 33,779
 ; REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 450-8400
 ; TELEFAX: (619) 587-5360
 ; INFORMATION FOR SEQ ID NO: 50:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2353 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: Internal
 ; ORIGINAL SOURCE:
 ; US-08-984-709A-50

Query Match 10.7%; Score 66.5; DB 4; Length 2353;
 Best Local Similarity 28.3%; Pred. No. 7.8;
 Matches 32; Conservative 12; Mismatches 40; Indels 29; Gaps 5;

QY 9 AILAFSLAOSFGAVCKEPEEVVPGGSRKRPDLQRLQRLFKSHS-SLEGILKLSQA 67
 DB 1977 ASLQIPLAVSPARSGEPHIALSPRG--TARSPSLRLLCROAVHTDSLEGRIDS----- 2030
 QY 68 SDDPKSTSPKRDHDFVGLMKRSVQDPDPTDVNOENVPSGLTKPPRA 120
 DB 2031 ---PRDTLDP-----ABEGEXTVPR--VYGGSLQSPPRS 2061

RESULT 13
 5268359-5

; Patent No. 5268359
 ; APPLICANT: HARMAR, ANTHONY J.; PASCALL, JOHN; MCKEOWN, ANN
 ; TITLE OF INVENTION: HUMAN TACHYKININS AND THEIR PRECURSOR
 ; NUMBER OF SEQUENCES: 7
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/285,964
 ; FILING DATE: 03-JUN-1987
 ; SEQ ID NO: 5;
 ; LENGTH: 126
 ; 5268359-5

Query Match 10.6%; Score 65.5; DB 6; Length 126;
 Best Local Similarity 38.2%; Pred. No. 1.6;
 Matches 26; Conservative 3; Mismatches 12; Indels 27; Gaps 4;

QY 34 GGRKRDPLQRLQRLFKSHSLE---GLTKAL---SOASTDPKESTSPKRDHDFV 87
 DB 66 GLMKRDLAD-----SSIEKQVALLKALYHGQIS-----HRRHKRDSRV 104
 QY 88 GLMKRSV 95
 DB 105 GLMKRAL 112

RESULT 14
 US-08-770-761A-8

; Sequence 8, Application US/08770761A
 ; Patent No. 5814503

; GENERAL INFORMATION:
 ; APPLICANT: Kovacevic, Steven
 ; APPLICANT: Rao, Ramachandra N.
 ; TITLE OF INVENTION: FUSION PROTEINS COMPRISING CELL CYCLE
 ; TITLE OF INVENTION: REGULATORY PROTEINS
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Eli Lilly and Company
 ; STREET: Lilly Corporate Center/Patent Division
 ; CITY: Indianapolis
 ; STATE: IN
 ; COUNTRY: USA
 ; ZIP: 46285

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/770,761A
 ; FILING DATE: 19-DEC-1996
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gaylo, Paul J.
 ; REGISTRATION NUMBER: 36,808
 ; REFERENCE/DOCKET NUMBER: X-10136
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 317-376-0756

TELEFAX: 317-277-1917
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 647 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-770-761A-8

Query Match 10.4%; Score 64.5; DB 2; Length 647;
Best Local Similarity 22.9%; Pred. No. 22;
Matches 30; Conservative 19; Mismatches 37; Indels 45; Gaps 7;

QY 16 AQSFGAVC-----KEPQEEVYPG-----GRSKRPD-----LYQLQRLFK--- 52
Db 182 AQTFFALCATDYKFLSNPPSVAAGSVAAVQGLNLRSPNNFLSYRLTRFLSRVAKCDP 241
QY 53 -----SHSLLEGILKA---LSQASTDPKESTSPKRDHDFVGLMGKRSVQPD---SPT 101
Db 242 DCLRACQEQIEALLESSLRQAQNNMDPKAAEEEEEEEE-----EVDLACTPT 290
QY 102 DVNOENVPSFG 112
Db 291 DVRDVDIASMG 301

RESULT 15
US-08-770-761A-7
: Sequence 7, Application US/0870761A
: Patent No. 5814503
: GENERAL INFORMATION:
: APPLICANT: Kovacevic, Steven
: APPLICANT: Otto, Keith A.
: APPLICANT: Rao, Ramachandra N.
: TITLE OF INVENTION: FUSION PROTEINS COMPRISING CELL CYCLE
: TITLE OF INVENTION: REGULATORY PROTEINS
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Eli Lilly and Company
: STREET: Lilly Corporate Center/Patent Division
: CITY: Indianapolis
: STATE: IN
: COUNTRY: USA
: ZIP: 46285
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/770,761A
: FILING DATE: 19-DEC-1996
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Gaylo, Paul J.
: REGISTRATION NUMBER: 36,808
: REFERENCE/DOCKET NUMBER: X-10136
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-376-0756
: TELEFAX: 317-277-1917
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 705 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-770-761A-7

Query Match 10.4%; Score 64.5; DB 2; Length 705;

Best Local Similarity 22.9%; Pred. No. 25;
Matches 30; Conservative 19; Mismatches 37; Indels 45; Gaps 7;

QY 16 AQSFGAVC-----KEPQEEVYPG-----GRSKRPD-----LYQLQRLFK--- 52
Db 228 AQTFFALCATDYKFLSNPPSVAAGSVAAVQGLNLRSPNNFLSYRLTRFLSRVAKCDP 287
QY 53 -----SHSLLEGILKA---LSQASTDPKESTSPKRDHDFVGLMGKRSVQPD---SPT 101
Db 288 DCLRACQEQIEALLESSLRQAQNNMDPKAAEEEEEEEE-----EVDLACTPT 336
QY 102 DVNOENVPSFG 112
Db 337 DVRDVDIASMG 347

Search completed: May 24, 2002, 17:04:21
Job time: 240 sec

PR 06-JUN-1997; 97US-0048970.
 PR 05-SEP-1997; 97US-0057765.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ferris AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;
 PI Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;
 PI Wei YF, Young PE, Zeng Z;
 XX
 DR WPI: 1998-520811/44.
 DR N-PSDB; AAV34302.

PT Isolated human poly:nucleotide(s) encoding secretory peptide(s) -
 PT used to develop products for the diagnosis and treatment of e.g.
 PT inflammation, cancers, CNS disorders or immune system disorders
 XX

PS Claim 1; Page 162-163; 201pp; English.

XX This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AAV34277) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 28 novel genes and their fragments (nucleic
 CC acid sequences: AAV34286-V34325; amino acid sequences AAV75196-W75235)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 28
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV34286 for described uses).
 XX

Sequence 121 AA:

Query Match 100.0%; Score 619; DB 19; Length 121;
 Best Local Similarity 100.0%; Pred. No. 2.6e-65;

Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLMLFTAILAFSLAOSFGAVCKEPEVEVPGGSRKRDPDLYQLQRLFKSHSLEGL 60
 DB 1 mrlmlftailafslagsfgavckepgeevpggskrdpdllyqlrlfkshslegl 60
 QY 61 LKALSGASTDPKESSTPEKRDHDFVGLMGKRSVQPSPTDVNGENVPSFGILKYPRA 120
 DB 61 lkalsgastdpkesstpekrdhdfvglmgrsvqpsptdvngenvpsfgilkypra 120
 QY 121 E 121
 DB 121 e 121

RESULT 2

AAW97213
 ID AAW97213 standard; Protein; 121 AA.

AC AAW97213;

DT 07-MAY-1999 (first entry)

DE A human zneurok1 polypeptide.

XX Human; zneurok1; neurokinin B: prohormone convertase; cell growth;
 XX modulation; inflammation; nociception; emesis; contraction;
 XX hormone secretion; DNA synthesis; inositol phosphate turnover;
 XX arachidonate release; phospholipase-C activation; gastric emptying;
 XX human neutrophil activation; ADCC capability;
 XX superoxide anion production; gene therapy.

OS Homo sapiens.
 XX
 XX W09855612-A1.
 PJ

XX
 PD 10-DEC-1998.
 XX

PF 28-MAY-1998; 98MO-US10842.

PR 02-JUN-1997; 97US-0048290.

PA (ZYMO) ZYMOGENETICS INC.

PI Sheppard PO;

DR WPI: 1999-070268/06.

DR N-PSDB; AAX15447.

PT New isolated neurokinin polypeptides, zneurok1 - used to develop
 PT products for modulating e.g. inflammation, nociception, emesis,
 PT muscle contraction, hormone secretion, DNA synthesis or cell growth
 XX

PS Claim 3; Page 76-77; 100pp; English.

XX The present sequence represents a human zneurok1 polypeptide. The
 CC polypeptide releases a neurokinin B polypeptide in the presence
 CC of a prohormone convertase capable of cleaving dibasic amino
 CC acids. The zneurok1 polypeptides can be used for modulating inflammation,
 CC agonists, antagonists or antibodies may also modulate inflammation,
 CC hormone secretion, DNA synthesis or cell growth, inositol phosphate
 CC turnover, arachidonate release, phospholipase-C activation, gastric
 CC emptying, human neutrophil activation or ADCC capability, or superoxide
 CC anion production. The polynucleotides can also be used for gene therapy.
 CC The products can also be used for detection, diagnosis and screening
 XX assays.

Sequence 121 AA:

Query Match 100.0%; Score 619; DB 20; Length 121;
 Best Local Similarity 100.0%; Pred. No. 2.6e-65;

Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLMLFTAILAFSLAOSFGAVCKEPEVEVPGGSRKRDPDLYQLQRLFKSHSLEGL 60
 DB 1 mrlmlftailafslagsfgavckepgeevpggskrdpdllyqlrlfkshslegl 60
 QY 61 LKALSGASTDPKESSTPEKRDHDFVGLMGKRSVQPSPTDVNGENVPSFGILKYPRA 120
 DB 61 lkalsgastdpkesstpekrdhdfvglmgrsvqpsptdvngenvpsfgilkypra 120
 QY 121 E 121
 DB 121 e 121

RESULT 3

AAW74413
 ID AAW74413 standard; Protein; 121 AA.

AC AAW74413;

DT 10-MAY-1999 (first entry)

DE HPMB091 protein sequence.

XX HPMB091; neurokinin B precursor; neuronal disorder; CNS related disorder;
 XX gastrointestinal disorder; cardiovascular disorder; metabolic disorder;
 XX smooth muscle disorder; inflammatory disorder; cancer; gene therapy.

OS Homo sapiens.

PN EP892053-A2.

PD 20-JAN-1999.

26-JUN-1998; 98EP-0305066.
14-JUL-1997; 97EP-0305215.
(HUMA-) HUMAN GENOME SCI.
(SMIK) SMITHKLINE BEECHAM PLC.
Duckworth DM, Hastings GA, Ruben SM;
WPI: 1999-083570/08.
N-PSDB: AAX18197.
New human neurokinin B precursor (HPMBQ91) polypeptides and polynucleotide - useful as diagnostic reagents and for prevention and treatment of neuronal, metabolic, inflammatory and gastrointestinal disorders, and cancers
Claim 11; Page 14-15; 18pp; English.

This sequence is the human neurokinin B precursor HPMBQ91 of the invention. HPMBQ91 polypeptides and polynucleotides are useful for diagnosing susceptibility to diseases associated with HPMBQ91 protein imbalance by determining HPMBQ91 polypeptide expression levels. HPMBQ91 polypeptides can be used to screen for agonists and antagonists by measuring the binding to HPMBQ91, and observing the stimulation or inhibition of HPMBQ91 function. These can be used in treatment to activate or inhibit HPMBQ91 activity to treat conditions associated with a lack of HPMBQ91 protein. Gene therapy may also be used to affect endogenous HPMBQ91 polypeptide expression. HPMBQ91 antibodies are useful for inducing an immune response to immunise and prevent disease, and for isolating HPMBQ91 clones or purifying the polypeptides by affinity chromatography. HPMBQ91 polypeptides can be administered directly or as a vaccine to inoculate against disease. Diseases diagnosed, prevented and treated include: neuronal disorders; CNS related disorders; gastrointestinal and cardiovascular disorders; metabolic disorders including diabetes and obesity; smooth muscle disorders; inflammatory disorders; and cancers including adenomas, leiomyomas, liposarcomas, melanomas, pulmonary chondroid hamartomas, lung, prostate and breast cancer. The HPMBQ91 polypeptide is also useful for mapping the gene to a chromosome, allowing gene inheritance to be studied through linkage analysis, and tissue localisation studies, for determining HPMBQ91 expression patterns.

Sequence 121 AA:

Query Match 100.0%; Score 619; DB 20; Length 121;
Best Local Similarity 100.0%; Pred. No. 2,6e-65;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRIMLFTAILAFSLAQSFGAVCKEPOEEVVPGGGRSKRDPDLYOLLQRLFKSHSSLEGL 60
|
1 mrimlftailafslasfgavckepoeevvp9ggrskrdpdllyqlqlfkshsslegl 60
61 LKALISQASTDPKKESTSPKRDMDHDFVGLMGKRSVQPSDSTVDNQENVSPFGILKYPPRA 120
|
61 lkalsgastdpkkestspkrdmdhdfvlgmgkrsvqpsdstvdnqenvspfgilkyppra 120
121 E 121
121 e 121

RESULT 4
AAB82380
ID AAB82380 standard; protein; 121 AA.

AC AAB82380;
XX 23-JUL-2001 (first entry)
XX Human neurokinin B precursor.
DE
XX

Neurokinin B; human; pregnancy; hypertension; pre-eclampsia;
diagnosis; therapy.
Homo sapiens.

Key Location/Qualifiers
FT Peptide 81..91
FT Peptide /label= Neurokinin
W0200136979-A2.

25-MAY-2001.

10-NOV-2000; 2000WO-GB04315.

16-NOV-1999; 99GB-0027125.

(UTRE-) UNITV READING.

Page N, Lowry P;

WPI: 2001-355676/37.
N-PSDB: AAF90333, AAF90334.

Detecting production of the human precursor of neurokinin B by the placenta in pregnancy induced hypertension or pre-eclampsia or related foetal complication

Example 1; Fig 1; 63pp; English.

The present sequence is that of human neurokinin B (NKB) precursor. The cloning of placental cDNA (see AAF90333) was used to identify the NKB precursor. The precursor is processed to the 10-amino acid NKB peptide. Detection of raised plasma levels of NKB, NKB precursor, its breakdown product or variants at an early stage of pregnancy provide an indication of the likely development of pregnancy induced hypertension or pre-eclampsia. Reduction in the levels of circulating NKB (or reduction of its effects) will ameliorate the adverse effects upon the mother seen in these conditions. Thus, the invention provides methods for predicting or diagnosing pregnancy induced hypertension, pre-eclampsia or related foetal complication based on measuring NKB levels in the blood, and methods for preventing or treating these conditions, e.g. by administering an agent that inhibits the biological effect of NKB, such as an NK1, NK2 or NK3 antagonist.

Sequence 121 AA:

Query Match 100.0%; Score 619; DB 22; Length 121;
Best Local Similarity 100.0%; Pred. No. 2,6e-65;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRIMLFTAILAFSLAQSFGAVCKEPOEEVVPGGGRSKRDPDLYOLLQRLFKSHSSLEGL 60
|
1 mrimlftailafslasfgavckepoeevvp9ggrskrdpdllyqlqlfkshsslegl 60
61 LKALISQASTDPKKESTSPKRDMDHDFVGLMGKRSVQPSDSTVDNQENVSPFGILKYPPRA 120
|
61 lkalsgastdpkkestspkrdmdhdfvlgmgkrsvqpsdstvdnqenvspfgilkyppra 120
121 E 121
121 e 121

RESULT 5
AB11878
ID AB11878 standard; peptide; 137 AA.

AC AB11878;
XX 11-JAN-2002 (first entry)
XX

DE XX Human neurokinin B-like protein homologue, SEQ ID NO:2248.
 XX XX
 KW KW Human: cytokine; cell proliferation; cell differentiation; growth factor;
 KW KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW KW chronic inflammatory condition; proliferative retinopathy;
 KW KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW KW bone disorder; osteoporosis; vascular growth disorder;
 KW KW tissue regeneration; wound healing; infection; immune disorder;
 KW KW cell culture; drug screening; gene therapy; antiinflammation;
 KW KW antiaesthetic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW KW antifungal; vulnery; antiulcer.
 OS XX Homo sapiens.
 XX XX WO200157188-A2.
 XX XX 09-AUG-2001.
 XX- XX- 05-FEB-2001; 2001WO-US03800.
 XX XX 03-FEB-2000; 2000US-0496914.
 PR XX 27-APR-2000; 2000US-0560875.
 XX- XX- (HXSE-) HXSEQ INC.
 PA XX
 XX XX Tang YT, Liu C, Drmanac RT;
 XX XX WPI: 2001-457740/49.
 DR XX N-PSDB: ABA09122.
 XX XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 XX e.g. arthritis and cancer -
 XX
 PS Claim 20; Page 270; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, and hence
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells

CC that can be used to augment or replace cells damaged by illness.
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 XX
 SQ Sequence 137 AA;
 XX
 Query Match 100.0%; Score 619; DB 22; Length 137;
 Best Local Similarity 100.0%; Pred. No. 3e-65;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRMILFTAILAFSLAQSGAVCKEPQEEVPGGGRKRDPLYOLGRFSSLSGL 60
 DB 17 mrmilftailafslagsfgavckepqeevpgggrskrdplyqlqlfshslegl 76
 QY 61 LKALSGASDPKRSSTPEKRDMDFFVGLMGRKRSVQPSPTDVNOENVPSEGLKYPRA 120
 DB 77 lkalsgastdpkrsstpekrdmdffvglmgrksvqpsptdvnngenvpsfgllkyppra 136
 QY 121 E 121
 DB 137 e 137
 XX
 RESULT 6
 AAW75228
 ID AAW75228 standard; Protein, 121 AA.
 XX
 AAW75228;
 AC XX
 XX XX
 DT 29-JAN-1999 (first entry)
 XX XX
 DE Human secreted protein encoded by gene 17 clone HPMB91.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW endocrine; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 67
 FT /label= unknown
 FT MISC-difference 89
 FT /label= unknown
 XX
 PN MO9840483-A2.
 XX
 PD 17-SEP-1998.
 XX
 PF 12-MAR-1998; 98WO-US04858.
 XX
 PR 19-DEC-1997; 97US-0068368.
 PR 14-MAR-1997; 97US-0040710.
 PR 14-MAR-1997; 97US-0040762.
 PR 30-MAY-1997; 97US-0048100.
 PR 30-MAY-1997; 97US-0048189.
 PR 30-MAY-1997; 97US-0048357.
 PR 30-MAY-1997; 97US-0050934.
 PR 06-JUN-1997; 97US-0048570.
 PR 05-SEP-1997; 97US-0057765.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ferris AM, Fischer CL, Gentz RL, Greene JM, Kyaw H,
 LI H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;

PI Wei YF, Young PE, Zeng Z;
 XX
 DR WPI: 1998-520811/44.
 DR N-PSDB: AAV34318.
 XX
 PT Isolated human poly(nucleotide(s) encoding secretory peptide(s) -
 PT used to develop products for the diagnosis and treatment of e.g.
 PT inflammation, cancers, CNS disorders or immune system disorders
 XX
 PS Claim 1; Page 175; 201pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AAV34277) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 28 novel genes and their fragments (nucleic
 CC acid sequences: AAV34286-V34325; amino acid sequences AAV75196-W75235)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 28
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV34286 for described uses).
 CC
 SQ Sequence 121 AA:

Query Match 98.5%; Score 610; DB 19; Length 121;
 Best Local Similarity 98.3%; Pred. No. 3e-64;
 Matches 119; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRLMLFTLILAFSLAQSGAVCKEPEVEVPGGGRSKRDPDLYQLQLRFLKSHSSLEGL 60
 Db 1 mrlmlftlilafslagsfgavckepqeevpgggrskrdpdllyqlqlrflshsslegl 60
 OY 61 LKALSOASDPKPESTSPERKDMHDFVGLMGKRSVOPDSPTDVNOENVPSFGILKYPRA 120
 Db 61 lkalsqstcdpkestspkrkdmhdfvglmgrsvopdsptdngenvpsfgilkyppra 120
 OY 121 E 121
 Db 121 e 121

RESULT 7
 AAM96144
 ID AAM96144 standard; Protein; 122 AA.
 XX
 AC AAM96144;
 XX
 DT 27-APR-1999 (first entry)
 XX
 DE Human preprotachykinin B.
 XX
 XX Preprotachykinin B, PPT-B; neoplastic disorder;
 KW neurological disorder; Alzheimer's disease; amnesia;
 KW cerebral neoplasms; dementia; depression; Down's syndrome;
 KW Huntington's disease; multiple sclerosis; Parkinson's disease;
 KW paranoid psychoses; schizophrenia; Tourette's disorder; angina;
 KW anaphylactic shock; asthma; cardiovascular shock;
 KW myocardial infarction; migraine.
 KW
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 104 /label= leu, Ser or Tyr
 XX
 PN W09857986-A2.
 XX
 PD 23-DEC-1998.

XX
 PF 19-JUN-1998; 98WO-US12855.
 XX
 PR 19-JUN-1997; 97US-0879995.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Hillman JT, Kaser MR, Lal P;
 XX
 DR WPI: 1999-080948/07.
 DR N-PSDB: AAX08906.
 XX
 CC New human preprotachykinin B - useful for treating neurological
 CC disorders and cancer
 PT
 PS Claim 1; Page 48-49; 57pp; English.
 XX
 CC Human preprotachykinin B (PPT-B) can be used to treat a
 CC neurological disorder. Antagonists of PPT-B can also be used in
 CC the treatment of neoplastic disorders. Particular neurological,
 CC disorders include akathisia, Alzheimer's disease, amnesia,
 CC amyotrophic lateral sclerosis, bipolar disorder, cataplexy,
 CC cerebral neoplasms, dementia, depression, Down's syndrome, tardive
 CC dyskinesia, dystonia, Huntington's disease, multiple sclerosis,
 CC neurofibromatosis, Parkinson's disease, paranoid psychoses,
 CC schizophrenia, and Tourette's disorder. PPT-B or its agonist can
 CC also be used to treat angina, anaphylactic shock, arrhythmias,
 CC asthma, cardiovascular shock, Cushing's syndrome, hypertension,
 CC hypoglycemia, myocardial infarction, migraine and pheochromocytoma.
 XX
 SQ Sequence 122 AA:

Query Match 97.2%; Score 601.5; DB 20; Length 122;
 Best Local Similarity 97.5%; Pred. No. 3.1e-63;
 Matches 119; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

OY 1 MRLMLFTLILAFSLAQSGAVCKEPEVEVPGGGRSKRDPDLYQLQLRFLKSHSSLEGL 60
 Db 1 mrlmlftlilafslagsfgavckepqeevpgggrskrdpdllyqlqlrflshsslegl 60
 OY 61 LKALSOASDPKPESTSPERKDMHDFVGLMGKRSVOPDSPTDVNOENVPSFGILKYPRA 119
 Db 61 lkalsqstcdpkestspkrkdmhdfvglmgrsvopdsptdngenvpsfgilkyppr 120
 OY 120 AE 121
 Db 121 ae 122

RESULT 8
 AAB33445
 ID AAB33445 standard; Protein; 135 AA.
 XX
 AC AAB33445;
 XX
 DT 29-JAN-2001 (first entry)
 XX
 DE Human PRO1155 protein UNQ585 SEQ ID NO:157.
 XX
 KW Human; immune related disease; diagnosis; antiinflammatory; cardiant;
 KW dermatological; antiarthritic; antiinflammatory; immunosuppressive;
 KW haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;
 KW antianemic; hepatotropic; virucide; antipsoriatic; antiallergic;
 KW antislumatic; systemic lupus erythematosus; rheumatoid arthritis;
 KW osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;
 KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
 KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
 KW autoimmune thrombocytopenia; immune mediated renal disease;
 KW demyelinating disease; hepatobiliary disease; Whipple's disease;
 KW inflammatory bowel disease; gluten-sensitive enteropathy;
 KW autoimmune disease; immune-mediated skin disease; allergic disease;
 KW immunological disease; transplantation associated disease;

KM graft rejection; graft-versus-host-disease.
 XX Homo sapiens.
 XX WO200053758-A2.
 XX
 XX 14-SEP-2000.
 PD
 XX
 PF 02-MAR-2000; 2000WO-US05841.
 XX
 PR 08-MAR-1999; 99WO-US05028.
 PR 10-MAR-1999; 99US-0123618.
 PR 12-MAR-1999; 99US-0123957.
 PR 23-MAR-1999; 99US-0125775.
 PR 12-APR-1999; 99US-0128849.
 PR 20-APR-1999; 99WO-US08615.
 PR 28-APR-1999; 99US-0131445.
 PR 04-MAY-1999; 99US-0132371.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 28-JUL-1999; 99US-0146222.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-OCT-1999; 99US-0162506.
 PR 29-NOV-1999; 99WO-US28214.
 PR 30-NOV-1999; 99WO-US28313.
 PR 30-NOV-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30999.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 XX
 PA (GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;
 PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;
 PI Stewart TA, Tumes D, Watanabe CK, Wood WL, Yan W;
 XX
 XX WPI; 2000-572271/53.
 DR N-PSDB; AAC58610.
 XX
 PT Sixty four PRO polypeptides, useful in the diagnosis and treatment of
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
 XX
 XX Claim 33; Fig 64; 309pp; English.
 XX
 CC The present invention describes sixty four human PRO proteins which can
 CC be used in the treatment of immune related diseases. The human PRO
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
 CC treating and diagnosing immune related disorders. The disorders are
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic

CC anaemia, autoimmune thrombocytopaenia, thyroiditis, diabetes mellitus,
 CC immune-mediated renal disease, demyelinating diseases of the central
 CC and peripheral nervous systems, hepatobiliary diseases, inflammatory
 CC bowel disease, gluten-sensitive enteropathy and Whipple's disease,
 CC autoimmune or immune-mediated skin diseases, allergic diseases,
 CC immunological diseases of the lung, and transplantation associated
 CC diseases including graft rejection and graft-versus-host-disease.
 CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and
 CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.
 XX
 SQ Sequence 135 AA:
 QY Query Match 80.1%; Score 496; DB 21; Length 135;
 Best local similarity 97.0%; Pred. No. 1, 1e-50;
 Matches 98; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 DB 1 MRLMLFTAILAFSLAQSRGAVCKEPEVEYGGGSKRDPDLYQLQRLFRSHSLEGL 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1 mrlmlftailafslagsigavckepqevvpqggrskdpdlyqlqlfkshslegl 60
 QY 61 LKALSGASTDPKESTSPKRDMDPEFVGLMGKRSVQPSPT 101
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 lkalsgastdpkestspkrdmdpfvglmgrsvqpspt 101
 DB 61 lkalsgastdpkestspkrdmdpfvglmgrsvqpspt 101
 RESULT 9
 AAY66739
 ID AAY66739 standard; protein; 135 AA.
 XX
 AC AAY66739;
 XX
 DT 05-APR-2000 (first entry)
 XX
 DE Membrane-bound protein PRO1155.
 XX
 DE Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
 XX
 KW pharmaceutical; receptor immunoadhesin; gene mapping.
 KW
 XX Homo sapiens.
 OS
 XX WO9963088-A2.
 PN
 XX
 PD 09-DEC-1999.
 XX
 XX 02-JUN-1999; 99WO-US12252.
 PF
 XX
 PR 02-JUN-1998; 98US-0087607.
 PR 02-JUN-1998; 98US-0087609.
 PR 03-JUN-1998; 98US-0087759.
 PR 03-JUN-1998; 98US-0087827.
 PR 04-JUN-1998; 98US-0088021.
 PR 04-JUN-1998; 98US-0088028.
 PR 04-JUN-1998; 98US-0088029.
 PR 04-JUN-1998; 98US-0088030.
 PR 04-JUN-1998; 98US-0088032.
 PR 04-JUN-1998; 98US-0088326.
 PR 05-JUN-1998; 98US-0088167.
 PR 05-JUN-1998; 98US-0088202.
 PR 05-JUN-1998; 98US-0088212.
 PR 05-JUN-1998; 98US-0088217.
 PR 05-JUN-1998; 98US-0088217.
 PR 09-JUN-1998; 98US-0088655.
 PR 10-JUN-1998; 98US-0088722.
 PR 10-JUN-1998; 98US-0088730.
 PR 10-JUN-1998; 98US-0088734.
 PR 10-JUN-1998; 98US-0088738.
 PR 10-JUN-1998; 98US-0088740.
 PR 10-JUN-1998; 98US-0088741.
 PR 10-JUN-1998; 98US-0088742.
 PR 10-JUN-1998; 98US-0088810.

```

PR 10-JUN-1998; 98US-0086811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090255.
PR 23-JUN-1998; 98US-0090259.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.

```

```

PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 11-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098525.
PR 16-SEP-1998; 98US-0100634.
PR 12-JAN-1999; 98US-0115565.

```

(GETH) GENENTECH INC.

Baker K, Chen J, Goddard A, Gurney AL, Smith V, Matanabe CK,
Wood WT, Yuan J;

WPI; 2000-072883/06.
N-PSDB; AAZ65085.

Membrane-bound proteins and related nucleotide sequences

claim 12; Fig 254; 822pp; English.

The invention provides membrane-bound PRO polypeptides and polynucleotides encoding them. The PRO sequences of the invention were identified based on extracellular domain homology screening. The PRO sequences have homology with proteins including LDL receptors, TIE ligands and various enzymes. The membrane-bound proteins and receptor molecules are useful as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be used as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. The PRO encoding sequences are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques.

Sequence 135 AA;

Query Match 80.1%; Score 496; DB 21; Length 135;
Best Local Similarity 97.0%; Pred. No. 1,1e-50;
Matches 98; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

QY      1 MRIMLFTAILAFSLAOSFGAVCKEPQEVYVGGGRSKRDPDLYQLRLFKSHSLEGL 60
Db      1 mrimlftailafslafslagsgfaveckepqeevvp9ggrskrdpdyqlrlfkshslegl 60
QY      61 LKALSQASTDPKESTSPKRDMDHFFVGLMGKRSVQPSDPT 101
Db      61 lkalsqastdpkestspkrdmdhdfvglmgrksrvqpsgk 101

RESULT 10
AAU29245 standard; Protein: 135 AA.
ID      AAU29245;
AC      AAU29245;
XX      18-DEC-2001 (first entry)
XX      Human PRO polypeptide sequence #222.
DE      Human PRO polypeptide sequence #222.
XX      PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
XX      dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
XX      blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
XX      adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX      Homo sapiens.
XX      MO200168848-A2.
XX      20-SEP-2001.
XX      28-FEB-2001; 2001WO-US06520.
XX      01-MAR-2000; 2000WO-US05601.
XX      02-MAR-2000; 2000WO-US05841.
XX      03-MAR-2000; 2000US-187202P.
XX      06-MAR-2000; 2000US-186968P.
XX      14-MAR-2000; 2000US-189320P.
XX      14-MAR-2000; 2000US-189320P.
XX      15-MAR-2000; 2000WO-US06884.
XX      21-MAR-2000; 2000US-190828P.
XX      21-MAR-2000; 2000US-191007P.
XX      21-MAR-2000; 2000US-191048P.
XX      21-MAR-2000; 2000US-191314P.
XX      28-MAR-2000; 2000US-192655P.
XX      29-MAR-2000; 2000US-193032P.
XX      29-MAR-2000; 2000US-193053P.
XX      30-MAR-2000; 2000WO-US08439.
XX      04-APR-2000; 2000US-194449P.
XX      04-APR-2000; 2000US-194647P.
XX      11-APR-2000; 2000US-195975P.
XX      11-APR-2000; 2000US-196000P.
XX      11-APR-2000; 2000US-196187P.
XX      11-APR-2000; 2000US-196590P.
XX      11-APR-2000; 2000US-196820P.
XX      18-APR-2000; 2000US-198121P.
XX      18-APR-2000; 2000US-198585P.
XX      25-APR-2000; 2000US-199397P.
XX      25-APR-2000; 2000US-199550P.
XX      25-APR-2000; 2000US-199654P.
XX      03-MAY-2000; 2000US-201516P.
XX      17-MAY-2000; 2000WO-US13705.
XX      22-MAY-2000; 2000WO-US14042.
XX      30-MAY-2000; 2000WO-US14941.
XX      02-JUN-2000; 2000WO-US15264.
XX      05-JUN-2000; 2000US-209832P.
XX      28-JUL-2000; 2000WO-US20710.
XX      22-AUG-2000; 2000US-064484B.
XX      24-AUG-2000; 2000WO-US23328.
XX      08-NOV-2000; 2000WO-US30952.
XX      01-DEC-2000; 2000WO-US32678.
XX      20-DEC-2000; 2000WO-US34956.
XX

```

```

PA      (GETH ) GENENTECH INC.
XX      Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
XX      Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX      WPI: 2001-602746/68.
XX      N-PSDB; AAS46146.
XX      Novel nucleic acids encoding PRO polypeptides, used to diagnose the
XX      presence of tumours, such as prostate and breast tumours, in mammals and
XX      to screen for modulators of the compounds -
XX      Claim 11; Fig 444; 774pp; English.
XX      Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
XX      The PRO polypeptides and their associated nucleic acids can be used to
XX      detect the presence of a tumour in a mammal by comparing the level of
XX      expression of a PRO polypeptide in a test sample of cells from the animal
XX      and a control sample of normal cells, whereby a higher level of
XX      expression in the test sample indicates the presence of a tumour in the
XX      mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
XX      and rabbits but are preferably human. The polypeptides can be used to
XX      stimulate tumour necrosis factor (TNF) alpha release from human blood,
XX      when contacted with it. A specific polypeptide can be used to stimulate
XX      the proliferation or differentiation of chondrocyte cells. The PRO
XX      proteins can be used to determine the presence of tumours and also
XX      susceptibility to tumour development, particularly adrenal, lung, colon,
XX      breast, prostate, rectal, cervical, or liver tumours, in mammalian
XX      subjects. The oligonucleotide probes specific for the PRO nucleic acids
XX      can be used for genetic analysis of individuals with genetic disorders.
XX      Sequence 135 AA:
XX      SQ
XX
QY      1 MRIMLFTAILAFSLAOSFGAVCKEPQEVYVGGGRSKRDPDLYQLRLFKSHSLEGL 60
Db      1 mrimlftailafslafslagsgfaveckepqeevvp9ggrskrdpdyqlrlfkshslegl 60
QY      61 LKALSQASTDPKESTSPKRDMDHFFVGLMGKRSVQPSDPT 101
Db      61 lkalsqastdpkestspkrdmdhdfvglmgrksrvqpsgk 101

RESULT 11
AAB65262
ID      AAB65262 standard; Protein: 135 AA.
XX      AAB65262;
XX      02-APR-2001 (first entry)
XX      Human PRO1155 (UNQ585) protein sequence SEQ ID NO:359.
XX      Human; secreted and transmembrane protein; PRO; cytosolic;
XX      cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
XX      diagnostic assay.
XX      Homo sapiens.
XX      OS
XX      MO200073454-A1.
XX      07-DEC-2000.
XX      30-MAR-2000; 2000WO-US08439.
XX      02-JUN-1999; 99WO-US12252.
XX      23-JUN-1999; 99US-0141037.
XX      07-JUL-1999; 99US-0143048.
XX      20-JUL-1999; 99US-0144758.
XX

```

```

PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 17-AUG-1999; 99US-0149386.
PR 15-SEP-1999; 99WO-US21050.
PR 15-SEP-1999; 99WO-US21547.
PR 08-OCT-1999; 99US-0158663.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 16-DEC-1999; 99WO-US30099.
PR 20-DEC-1999; 99WO-US30919.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
XX
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
PI Ferrara N, Fong S, Gerber H, Gerlitsen ME, Goddard A, Godowski PJ,
PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF,
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI,
PI Zhang Z.
XX
XX WPI; 2001-032160/04.
DR N-PSDB; AAF44231.
XX
XX PRO polynucleotides used to produce polypeptides used to target
PT bioactive molecules such as toxins, radiolabels or antibodies, to
PT specific cells, to cause targeted cell death -
XX
XX Claim 12; Fig 254; 935pp; English.
XX
XX The present invention describes human secreted and transmembrane PRO
CC proteins. The PRO proteins have cytosolic activity. The PRO proteins
CC can be used for targeted delivery of bioactive molecules, such as
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC sequences, and their fragments, can be used as hybridization probes, in
CC chromosomal and gene mapping, and in the generation of anti-sense RNA
CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
XX Sequence 135 AA:
SQ

```

```

Query Match 80.1%; Score 496; DB 22; Length 135;
Best Local Similarity 97.0%; Pred. No. 1.1e-50;
Matches 99; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 MRIMLFTAIATFSLAOSFGAVCKEPEEYVPGGSKRBDPLDYQ---LQRLFKSH-S 60
DB 1 mrimlftaiaatafslafagvckepqevvpggskrbdpldyqlqrlfkshslegl 60
OY 61 LKALSOASTDPKESTSPKRDMDHDFVGLMGRSVQDPSPT 101
DB 61 lkalsqastdpkestspkrdmdhdfvglmgrsvqdpst 101

```

```

RESULT 12
AAW96145
ID AAW96145 standard; Protein: 126 AA.
XX

```

```

AC AAW96145;
XX
XX 27-APR-1999 (first entry)
DT
XX
XX Bovine preprotachykinin B.
DE
XX
XX Preprotachykinin B; PPT-B; neoplastic disorder;
KW neurological disorder; Alzheimer's disease; amnesia;
KW cerebral neoplasms; dementia; depression; Down's syndrome;
KW Huntington's disease; multiple sclerosis; Parkinson's disease;
KW paranoid psychoses; schizophrenia; Tourette's disorder; angina;
KW anaphylactic shock; asthma; cardiovascular shock;
KW myocardial infarction; migraine.
XX
XX Bos taurus.
OS
XX
XX WO9857986-A2.
PN
XX
XX 23-DEC-1998.
PD
XX
XX 19-JUN-1998; 98WO-US12855.
PF
XX
XX 19-JUN-1997; 97US-0879995.
PR
XX
XX (INCY-) INCYTE PHARM INC.
PA
XX
XX Hillman JT, Kaser MR, Lal P.
PI
XX
XX WPI; 1999-080948/07.
DR
XX
XX New human preprotachykinin B - useful for treating neurological
PT disorders and cancer
PT
XX
XX Disclosure; Figure 2; 57pp; English.
PS
XX
XX Human preprotachykinin B (AAW96144); (PPT-B) can be used to treat a
CC neurological disorder. Antagonists of PPT-B can also be used in
CC the treatment of neoplastic disorders. Particular neurological,
CC disorders include akathisia, Alzheimer's disease, catatonia,
CC amyotrophic lateral sclerosis, bipolar disorder, catatonia, tardive
CC cerebral neoplasms, dementia, depression, Down's syndrome, tardive
CC dyskinesia, dystonias, Huntington's disease, multiple sclerosis,
CC neurofibromatosis, Parkinson's disease, paranoid psychoses,
CC schizophrenia, and Tourette's disorder. PPT-B or its agonist can
CC also be used to treat angina, anaphylactic shock, arrhythmias,
CC asthma, cardiovascular shock, Cushing's syndrome, hypertension,
CC hypoglycemia, myocardial infarction, migraine and pheochromocytoma.
XX
XX Sequence 126 AA:
SQ

```

```

Query Match 67.3%; Score 416.5; DB 20; Length 126;
Best Local Similarity 65.9%; Pred. No. 2.6e-41;
Matches 83; Conservative 15; Mismatches 23; Indels 5; Gaps 2;

```

```

OY 1 MRIMLFTAIATFSLAOSFGAVCKEPEEYVPGGSKRBDPLDYQ---LQRLFKSH-S 55
DB 1 mrimlftaiaatafslafagvckepqevvpggskrbdpldyqlqrlfkshslegl 60
OY 56 SLGLKALSOASTDPKESTSPKRDMDHDFVGLMGRSVQDPSPTDVNOENVPSFGILK 115
DB 61 slgllkalsqastdpkestspkrdmdhdfvglmgrsvqdpstpvnoenvpsfgil 120
OY 116 YPPRAE 121
DB 121 yppsve 126

```

```

RESULT 13
AAW96146
ID AAW96146 standard; Protein: 116 AA.
XX
XX AAW96146;

```

```

XX 27-APR-1999 (first entry)
XX Rat preprotachykinin B.
DE
XX
XX Preprotachykinin B; PPT-B; neoplastic disorder;
XX neurological disorder; Alzheimer's disease; amnesia;
XX cerebral neoplasms; dementia; depression; Down's syndrome;
XX Huntington's disease; multiple sclerosis; Parkinson's disease;
XX paranoid psychoses; schizophrenia; Tourette's disorder; angina;
XX anaphylactic shock; asthma; cardiovascular shock;
XX myocardial infarction; migraine.
XX
XX Rattus rattus.
XX
XX WO9857986-A2.
XX
XX 23-DEC-1998.
XX
XX 19-JUN-1998; 98WO-US12855.
XX
XX 19-JUN-1997; 97US-0879995.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Hillman JJ, Kaser MR, Lal P;
XX WPI; 1999-080948/07.
XX
XX New human preprotachykinin B - useful for treating neurological
XX disorders and cancer
XX
XX PS Disclosure; Figure 2: 57pp; English.
XX
XX Human preprotachykinin B (AAM96144) (PPT-B) can be used to treat a
XX neurological disorder. Antagonists of PPT-B can also be used in
XX the treatment of neoplastic disorders. Particular neurological,
XX disorders include akathisia, Alzheimer's disease, amnesia,
XX anorectic lateral sclerosis, bipolar disorder, catatonia,
XX cerebral neoplasms, dementia, depression, Down's syndrome, tardive
XX dyskinesia, dystonias, Huntington's disease, multiple sclerosis,
XX neurofibromatosis, Parkinson's disease, paranoid psychoses,
XX schizophrenia, and Tourette's disorder. PPT-B or its agonist can
XX also be used to treat angina, anaphylactic shock, arrhythmias,
XX asthma, cardiovascular shock, Cushing's syndrome, hypertension,
XX hypoglycemia, myocardial infarction, migraine and pheochromocytoma.
XX
XX SO Sequence 116 AA:

Query Match 57.1%; Score 353.5; DB 20; Length 116;
Best Local Similarity 61.7%; Pred. No. 6.7e-34;
Matches 74; Conservative 16; Mismatches 21; Indels 9; Gaps 3;

OY 1 MRMLFTYIILAFSLAQSGAVCKEPOEEVYPGGGRSKRDPDLY-----QLLQRLFKSHS- 55
DB 1 mrsamlfaavialslawtfgaaceepgq-----ggrlskdsdlsjpppllrlydsrsi 56
OY 56 SLEGLIKALSOASTDPKESTPEKRDMDHFFVGLMGKRSVDPDSTDVNOENVPSFGILK 115
DB 57 slgellikvlskasvpgktslpqkrmdhffvglmgrksnspqdpdpadvveentpsfgylk 116

RESULT 14
AAM97214
ID AAM97214 standard; Protein: 92 AA.
XX
XX AAM97214;
XX
XX 07-MAY-1999 (first entry)
XX
XX A murine homologue of human zneurok1 polypeptide.
XX

```

```

KW Human; zneurok1; neurokinin B; prohormone convertase; cell growth;
KW modulation; inflammation; nonreception; emesis; contraction;
KW hormone secretion; DNA synthesis; inositol phosphate turnover;
KW arachidonate release; phospholipase-C activation; gastric emptying;
KW human neutrophil activation; ADCC capability;
KW superoxide anion production; gene therapy.
XX
XX Mus sp.
XX
XX WO9855612-A1.
XX
XX 10-DEC-1998.
XX
XX 28-MAY-1998; 98WO-US10842.
XX
XX 02-JUN-1997; 97US-0048290.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Sheppard PO;
XX
XX WPI; 1999-070268/06.
XX
XX N-PSDB; AAX15448.
XX
XX New isolated neurokinin polypeptides; zneurok1 - used to develop
XX products for modulating e.g. inflammation, nonreception, emesis,
XX muscle contraction, hormone secretion, DNA synthesis or cell growth
XX
XX Example 1; Page 87; 100pp; English.
XX
XX The present sequence represents a murine homologue of human zneurok1. The
XX zneurok1 polypeptide releases a neurokinin B polypeptide in the presence
XX of a prohormone convertase capable of cleaving dibasic amino
XX acids. The zneurok1 polypeptides can be used for modulating inflammation,
XX nonreception or emesis. The polypeptides, fragments, fusion proteins,
XX agonists, antagonists or antibodies may also modulate contraction,
XX hormone secretion, DNA synthesis or cell growth, inositol phosphate
XX turnover, arachidonate release, phospholipase-C activation, gastric
XX emptying, human neutrophil activation or ADCC capability, or superoxide
XX anion production. The polynucleotides can also be used for gene therapy.
XX The products can also be used for detection, diagnosis and screening
XX assays.
XX
XX SO Sequence 92 AA:

Query Match 43.7%; Score 270.5; DB 20; Length 92;
Best Local Similarity 61.5%; Pred. No. 3.3e-24;
Matches 59; Conservative 12; Mismatches 16; Indels 9; Gaps 3;

OY 1 MRIMLFTYIILAFSLAQSGAVCKEPOEEVYPGGGRSKRDPDLY-----LLQRLFKSHS- 55
DB 1 mrsamlfaavialslawtfgavcepgq-----ggrlskdsdlyqpsllrlydsrpy 56
OY 56 SLEGLIKALSOASTDPKESTPEKRDMDHFFVGLMG 91
DB 57 slgellikvlskasvpgktslpqkrmdhffvglmrg 92

RESULT 15
AA12634
ID AA12634 standard; Protein: 51 AA.
XX
XX AA12634;
XX
XX 22-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO: 299 from WO 9906553.
XX
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW

```

KW reproductive hormone regulation; chemotactic; chemokine; haemostatic;
 KW thrombolytic; antiinflammatory; tumour inhibition; antitumour.
 OS Homo sapiens.
 PN MO9906553-A2.
 PD 11-FEB-1999.
 PF 31-JUL-1998; 98WO-1B01237.
 PR 01-AUG-1997; 97US-0905051.
 PA (GEST) GENSET.
 PI Duclert A, Dumas Milne Edwards J, Lacroix B;
 DR MPI: 1999-153783/13.
 DR N-PSDB: MAX41492.
 PT New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from umbilical cord, lymph ganglia,
 PT lymphocytes and placental tissue
 PS Claim 34; Page 389; 411pp; English.
 CC AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAX12521 to
 CC AAX12668, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, antiinflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 SO Sequence 51 AA:

Query Match 39.9%; Score 247; DB 20; Length 51;
 Best Local Similarity 100.0%; Pred. No. 9,1e-22;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRIMLFTAIATFSLAQSGAVCKEPEEVVPGGSRKRDPDLVQLQR 49
 |||||||
 Db 1 mrimlftaiaiafslagsgavckepgeevvpggskrdpdlvqlqr 49

Search completed: May 24, 2002, 17:04:01
 Job time: 820 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 17:05:21 ; Search time 25.18 Seconds

(without alignments)
831.310 Million cell updates/sec

Title: US-09-852-659-68

Perfect score: 121

Sequence: 1 MRMLFTALFSLAOSFC.....DVNOENPSFGILKYPRAE 121

Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	12.4	116	11	O9QXS9
2	8	6.6	72	4	O9Y494
3	8	6.6	114	6	O97947
4	8	6.6	115	11	O920K1
5	8	6.6	128	4	O9Y6V5
6	8	6.6	129	6	O97948
7	8	6.6	130	11	O9Z0K0
8	8	6.6	812	16	O9RT53
9	7	5.8	50	8	O9BAP5
10	7	5.8	98	8	O9B2G1
11	7	5.8	143	16	P73929
12	7	5.8	216	17	O97527
13	7	5.8	286	5	O9YFRO
14	7	5.8	297	8	O03305
15	7	5.8	348	9	O9MCR7
16	7	5.8	348	9	O9MCR6

17	7	5.8	350	9	O9MCR8	O9mcb8 lactococcus
18	7	5.8	366	8	O9Z255	O9z255 neurospora
19	7	5.8	377	16	O9PH30	O9ph30 xylella fas
20	7	5.8	384	5	O9N484	O9n484 caenorhabdi
21	7	5.8	388	5	O44330	O44330 mandua sex
22	7	5.8	394	16	O9RVP4	O9rvp4 deinococcus
23	7	5.8	405	5	O9NFC7	O9nfc7 drosophila
24	7	5.8	418	16	O9CLM9	O9clm9 pasteurella
25	7	5.8	475	16	O9PPC5	O9ppc5 campylobact
26	7	5.8	480	7	O38300	O38300 lactococcus
27	7	5.8	481	9	O38239	O38239 lactococcus
28	7	5.8	483	10	O9HWT2	O9hwt2 brassica ju
29	7	5.8	486	10	O9ACR2	O9acr2 arabidopsis
30	7	5.8	490	16	O97SN9	O97sn9 streptococc
31	7	5.8	493	10	P92947	P92947 arabidopsis
32	7	5.8	493	10	O9CAK5	O9cak5 arabidopsis
33	7	5.8	496	16	O9HWJ2	O9hwj2 pseudomonas
34	7	5.8	502	12	O91PQ1	O91pq1 tt virus. o
35	7	5.8	507	4	O96K90	O96k90 homo sapien
36	7	5.8	533	2	O9Z9N5	O9z9n5 bacillus ha
37	7	5.8	534	16	O9KE54	O9ke54 bacillus ha
38	7	5.8	553	10	O94GB5	O94gb5 oryza sativ
39	7	5.8	608	5	O9Y6D4	O9y6d4 drosophila
40	7	5.8	619	17	O981C7	O981c7 sulfolobus
41	7	5.8	637	11	O9CZ01	O9cz01 mus musculu
42	7	5.8	651	16	O9CPW7	O9cpw7 pasteurella
43	7	5.8	654	10	O9ZV25	O9zv25 arabidopsis
44	7	5.8	702	10	O9FR11	O9fr11 arabidopsis
45	7	5.8	712	5	O76408	O76408 caenorhabdi

ALIGNMENTS

RESULT 1
ID O9QXS9 PRELIMINARY; PRT; 116 AA.
AC O9QXS9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)
DE NEUROKININ B-LIKE PROTEIN ZNEUROK1.
GN TAC2 OR ZNEUROK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBL_TaxID-10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Sheppard P., Jellinek L., Whitmore T., Blumberg H., Lehner J.,
RA O'Hara P.;
RT "Mus musculus homolog of neurokinin B."
RT Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF186116; AAF01434.1; -
DR MGI; MGI:96476; Tac2.
DR InterPro; IPR003635; Neurokinin.
DR InterPro; IPR002040; Tachykinin.
DR Prodom; PD020370; Neurokinin; 1.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN 1.
SQ SEQUENCE 116 AA; 12737 MW; BDACB8171C2213CC CRC64;

Query Match 12.4%; Score 15; DB 11; Length 116;
Best local similarity 100.0%; Pred. No. 4e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 KRDMHDFVGLMGR 93
DB 80 KRDMHDFVGLMGR 94

RESULT 2
O9Y494

ID Q9Y494 PRELIMINARY; PRT: 72 AA.
AC Q9Y494;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GAMMA PREPROTACHYKININ (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD, BRAIN;
RA Lai J.P., Douglas S.D., Rappaport E., Wu J.M., Ho W.Z.;
RT "Identification of a Delta Isoform of preprotachykinin mRNA in Human
RT Mononuclear Phagocytes and Lymphocytes."
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF050657; AAC15703.1; -
DR InterPro: IPR003580; Protachykinin.
DR InterPro: IPR002040; Tachykinin.
DR Pfam: PF02202; Tachykinin; 1.
DR ProDom: PD005598; Protachykinin; 1.
DR SMART; SM00203; TK; 2.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN_2.
FT NON_TER 1 1
FT CHAIN 72 72
SQ SEQUENCE 72 AA; 8274 MW; 2C02BBA41EAAD16 CRC64;

Query Match 6.6%; Score 8; DB 4; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93
Db 53 FVGLMGKR 60

RESULT 3
ID Q97947 PRELIMINARY; PRT: 114 AA.
AC Q97947;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE GAMMA PREPROTACHYKININ I.
OS Tupia belangeri (northern tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupia.
OX NCBI_TaxID=37347;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Heiland A., Maeger H.J., Krinoeff M., Forssmann W.G.;
RT "Tachykinin precursors are highly conserved among different mammals."
RT Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z50785; CAA90648.1; -
DR InterPro: IPR003580; Protachykinin.
DR InterPro: IPR002040; Tachykinin.
DR Pfam: PF02202; Tachykinin; 1.
DR ProDom: PD005598; Protachykinin; 1.
DR SMART; SM00203; TK; 2.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN_2.
FT CHAIN 58 68
FT CHAIN 72 92
FT CHAIN 83 92
FT CHAIN 114 114
SQ SEQUENCE 114 AA; 13281 MW; B439C3D27FDA7CAB CRC64;

Query Match 6.6%; Score 8; DB 6; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93
Db 88 FVGLMGKR 95

RESULT 4
ID Q9Z0K1 PRELIMINARY; PRT: 115 AA.
AC Q9Z0K1;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE GAMMA PREPROTACHYKININ I.
OS Cavia porcellus (guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PIRRIDGE WHITE; TISSUE=BRAIN;
RA Heiland A., Maeger H.J., Krinoeff M., Forssmann W.G.;
RT "Tachykinin precursors are highly conserved among different mammals."
RT Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z50783; CAA90646.1; -
DR InterPro: IPR003580; Protachykinin.
DR InterPro: IPR002040; Tachykinin.
DR ProDom: PD005598; Protachykinin; 1.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN_1.
FT CHAIN 58 68
FT CHAIN 72 92
FT CHAIN 83 92
FT CHAIN 115 115
SQ SEQUENCE 115 AA; 13190 MW; 39EFEB8CB8A7174 CRC64;

Query Match 6.6%; Score 8; DB 11; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93
Db 88 FVGLMGKR 95

RESULT 5
ID Q9Y6V5 PRELIMINARY; PRT: 128 AA.
AC Q9Y6V5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE WUGSC:H.DJ0841B21.1 PROTEIN.
GN WUGSC:H.DJ0841B21.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC "The sequence of Homo sapiens PAC clone RP5-841B21."
RT Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Waterston R.;
RA Heiland A., Maeger H.J., Krinoeff M., Forssmann W.G.;
RT "Tachykinin precursors are highly conserved among different mammals."
RT Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC004140; AAC02754.1; -
DR InterPro: IPR003580; Protachykinin.
DR InterPro: IPR002040; Tachykinin.
DR Pfam: PF02202; Tachykinin; 1.
DR ProDom: PD005598; Protachykinin; 1.
DR SMART; SM00203; TK; 1.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN_2.
SQ SEQUENCE 128 AA; 14770 MW; 0F8D61774AFC1CA CRC64;

Query Match 6.6%; Score 8; DB 4; Length 126;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 86 FVGLMGKR 93
 |||||
 DB 95 FVGLMGKR 102

RESULT 6
 097948 PRELIMINARY; PRT; 129 AA.

AC 097948: 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE BETA PREPROTACHYKININ I.
 OS Tupala belangeri (northern tree shrew).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Scandentia; Tupalidae; Tupala.
 NCBI_TaxID=37347;
 OX NCBI_TaxID=37347;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Heitland A., Maegert H.J., Kruboeffer M., Forssmann W.G.;
 RT "Tachykinin precursors are highly conserved among different mammals."
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z50786; CAA90649.1; Protachykinin.
 DR InterPro: IPR003580; Protachykinin.
 DR InterPro: IPR002040; Tachykinin.
 DR Pfam: PF02202; Tachykinin; 1.
 DR ProDom: PD005598; Protachykinin; 1.
 DR SMART: SM00203; TK; 2.
 DR PROSITE: PS00267; TACHYKININ; UNKNOWN_2.
 FT CHAIN 58 68 SUBSTANCE P.
 FT CHAIN 72 107 NEUROKININ A.
 FT CHAIN 98 107 NEUROKININ A.
 SQ SEQUENCE 129 AA; 14941 MW; 5855E7ADC2D8674E CRC64;

Query Match 6.6%; Score 8; DB 6; Length 129;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 86 FVGLMGKR 93
 |||||
 DB 103 FVGLMGKR 110

RESULT 7
 0920X0 PRELIMINARY; PRT; 130 AA.

AC 0920X0: 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE BETA PREPROTACHYKININ I.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Hystriognathl; Caviidae; Cavia.
 NCBI_TaxID=10141;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=PIRBRIDGE WHITE; TISSUE=BRAIN;
 RA Heitland A., Maegert H.J., Kruboeffer M., Forssmann W.G.;
 RT "Tachykinin precursors are highly conserved among different mammals."
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z50784; CAA90647.1; -
 DR InterPro: IPR003580; Protachykinin.
 DR InterPro: IPR002040; Tachykinin.
 DR ProDom: PD005598; Protachykinin; 1.

DR PROSITE: PS00267; TACHYKININ; UNKNOWN_1.
 FT CHAIN 58 68 SUBSTANCE P.
 FT CHAIN 72 107 NEUROPEPTIDE K.
 FT CHAIN 98 107 NEUROKININ A.
 SQ SEQUENCE 130 AA; 14850 MW; C4B2F55B6A0A7C0 CRC64;

Query Match 6.6%; Score 8; DB 11; Length 130;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 86 FVGLMGKR 93
 |||||
 DB 103 FVGLMGKR 110

RESULT 8
 09RT53 PRELIMINARY; PRT; 812 AA.

AC 09RT53: 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE DNA GYRASE, SUBUNIT A.
 GN DR1913.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI.
 RX MEDLINE=20036896; PubMed-10567266;
 RA Whittle O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dolsen R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Uterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1."
 RL Science 286:1571-1577(1999).
 DR EMBL: AE002030; AAF11467.1; -
 DR HSSP: P09097; IAB4.
 DR TIGR: DR1913; -
 DR InterPro: IPR002205; DNA_topoisom.;
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00521; DNA_topoisom.; 1.
 DR SMART: SM00434; TOPAC; 1.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 812 AA; 89824 MW; C6F8596AB57BEC00 CRC64;

Query Match 6.6%; Score 8; DB 16; Length 812;
 Best Local Similarity 100.0%; Pred. No. 8.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 LREGLKAL 64
 |||||
 DB 377 LREGLKAL 384

RESULT 9
 09BBP5 PRELIMINARY; PRT; 50 AA.

AC 09BBP5: 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
 DE RIBOSOMAL PROTEIN L32.
 GN RPL32.
 OS Lotus japonicus.

OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
 OX NCBI_TaxID=34305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ACCESSION MG-20;
 RA Kato T.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ACCESSION MG-20;
 RA MEDLINE-21082929; PubMed-11214967;
 RA Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.;
 RT "Complete structure of the chloroplast genome of a legume, Lotus japonicus."
 RL DNA Res. 7:323-330(2000).
 DR EMBL: AP002983; BAB33243.1;
 DR InterPro: IPR002677; Ribosomal_L32p.
 DR Pfam: PF01783; Ribosomal_L32p; 1.
 KW Chloroplast.
 SQ SEQUENCE 50 AA; 5844 MW; 5589DC533C9ECB6 CRC64;

Query Match 5.8%; Score 7; DB 8; Length 50;
 Best Local Similarity 100.0%; Pred. No. 7.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 12 AFSLAOS 18
 Db 31 AFSLAOS 37

RESULT 10
 ID 09B2G1 PRELIMINARY; PRT; 98 AA.
 AC 09B2G1;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE NADH DEHYDROGENASE SUBUNIT 4L.
 GN NADH4L.
 OS Isoodon macrourus (Short-nosed bandicoot).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Peramellemorphia; Peramelidae; Isoodon.
 OX NCBI_TaxID=37698;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Phillips M.J., Lin Y.-H., Harrison G.L., Penny D.;
 RT "Mitochondrial Genomes of a Bandicoot and a Brushtail Possum Confirm the Monophyly of Australidelphian Marsupials."
 RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lin Y.-H., Phillips M.J.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF358864; AAK38289.1;
 DR InterPro: IPR003214; Mlt_MNDHb_oxidredctse_4L.
 DR InterPro: IPR001133; Oxidored_g2.
 DR Pfam: PF00420; oxidored_g2; 1.
 DR Pfam: PF000359; Mlt_MNDHb_oxidredctse_4L; 1.
 KW Mitochondrion.
 SQ SEQUENCE 98 AA; 10770 MW; 69332BFA406689D4 CRC64;

Query Match 5.8%; Score 7; DB 8; Length 98;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 10 IIAFSIA 16
 Db 11 IIAFSIA 16

Db 9 IIAFSIA 15
 RESULT 11
 ID P73929 PRELIMINARY; PRT; 143 AA.
 AC P73929;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE HYPOTHELICAL 15.3 KDA PROTEIN.
 GN SLR2101.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97061201; PubMed-8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hikosawa M., Sugita M., Saito S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions."
 RL DNA Res. 3:109-136(1996).
 DR EMBL: D90910; BAAL7995.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 143 AA; 15349 MW; 67A2C0F7A4BFF15C CRC64;

Query Match 5.8%; Score 7; DB 16; Length 143;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 108 VPSFGIL 114
 Db 75 VPSFGIL 81

RESULT 12
 ID 097527 PRELIMINARY; PRT; 216 AA.
 AC 097527;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PUTATIVE URACIL PHOSPHORIBOSYLTRANSFERASE.
 GN ST0281.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JCM 10545 / 7;
 RX PubMed-11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankal A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermophilic Crenarchaeon, Sulfolobus tokodaii strain7."
 RL DNA Res. 8:123-140(2001).
 DR EMBL: AP000982; BAB5251.1;
 DR Transferase; Glycosyltransferase; Hypothetical protein;
 KW Complete proteome.
 SQ SEQUENCE 216 AA; 24181 MW; EB511906282D3282 CRC64;

Query Match 5.8%; Score 7; DB 17; Length 216;
 Db 17 IIAFSIA 16

Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 EGLKAL 64
|||||
Db 87 EGLKAL 93

RESULT 13
Q9VFR0 PRELIMINARY; PRT; 286 AA.
AC Q9VFR0: 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE CG9286 PROTEIN.
GN CG9286.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Besson K.Y., Bohns P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Bokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Neison D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003701; AAF54950.1;
DR FLYBase; FBgn0038183; CG9286.
SQ SEQUENCE 286 AA; 32212 MW; 4ADB4B95F2991CD9 CRC64;

QY 45 QLIQRL 51
|||||
Db 51 QLIQRL 57

RESULT 14
ID 003305 PRELIMINARY; PRT; 297 AA.
AC 003305;
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CYTOCHROME B (FRAGMENT).
GN CYTB.
OS Carettochelys insculpta (pitted-shelled turtle).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Testudines; Cryptodira; Trionychoidea; Carettochelyidae;
OC Carettochelys.
OX NCBI_TaxID=44489;
RN [1]
RP SEQUENCE FROM N.A.
RA Shaffer H.B., Meylan P., McKnight M.L.;
RT "Tests of turtle phylogeny: molecular, morphological, and
RT paleontological approaches."
RL Syst. Biol. 0:0-(1996).
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -1- COPROCTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; U81355; AB57646.1;
DR InterPro: IPR000179; Cyt_b_b6.
DR Pfam: PF00032; cytochrome_b_c1.
DR Pfam: PF00033; cytochrome_b_n; 1.
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
FT Transmembrane. 1
FT NON_TER 1
FT NON_TER 297
SQ SEQUENCE 297 AA; 33587 MW; EA3172420509EBIC CRC64;

Query Match 5.8%; Score 7; DB 8; Length 297;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IMLEFTA 9
|||||
Db 91 IMLEFTA 97

RESULT 15
ID 09MCB7 PRELIMINARY; PRT; 348 AA.
AC 09MCB7;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE MAJOR CAPSID PROTEIN.
OS Lactococcus lactis bacteriophage Q38.
OC Viruses.
OX NCBI_TaxID=100636;
RN [1]
RP SEQUENCE FROM N.A.
RA Labrie S., Moineau S.;
RT "Multiplex PCR method for the detection and the identification of
RT Lactococcal bacteriophages in cheddar cheese whey".
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF152411; AAF85637.1; 9FSC77962A/AA296 CRC64;
SQ SEQUENCE 348 AA: 38887 MW: 9FSC77962A/AA296 CRC64;

Query Match 5.8%; Score 7; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 AOSFGAV 22
DB 273 AOSFGAV 279

Search completed: May 24, 2002, 17:07:25
Job time: 124 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 17:05:36 ; Search time 11.87 Seconds

(Without alignments)
394.698 Million cell updates/sec

Title: US-09-852-659-68
Perfect score: 121

Sequence: 1 MRMMLFTALIAFLASLQSG.....DVNQENVPSGLIKYPPRAE 121

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size: 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	100.0	121	1	TRNK_HUMAN
2	15	12.4	116	1	TRNK_MOUSE
3	15	12.4	116	1	TRNK_MOUSE
4	15	12.4	116	1	TRNK_MOUSE
5	10	8.3	10	1	TRNK_MOUSE
6	9	7.4	319	1	SRPH_SYNP7
7	8	6.6	115	1	TRNK_RABIT
8	8	6.6	129	1	TRNK_MOUSE
9	8	6.6	130	1	TRNK_MOUSE
10	8	6.6	130	1	TRNK_MOUSE
11	8	6.6	130	1	TRNK_MOUSE
12	8	6.6	130	1	TRNK_MOUSE
13	8	6.6	130	1	TRNK_MOUSE
14	7	5.8	310	1	TRNK_MOUSE
15	7	5.8	590	1	TRNK_MOUSE
16	7	5.8	590	1	TRNK_MOUSE
17	7	5.8	637	1	TRNK_MOUSE
18	7	5.8	648	1	TRNK_MOUSE
19	7	5.8	792	1	TRNK_MOUSE
20	6	5.0	63	1	TRNK_MOUSE
21	6	5.0	73	1	TRNK_MOUSE
22	6	5.0	73	1	TRNK_MOUSE
23	6	5.0	77	1	TRNK_MOUSE
24	6	5.0	104	1	TRNK_MOUSE
25	6	5.0	108	1	TRNK_MOUSE
26	6	5.0	108	1	TRNK_MOUSE
27	6	5.0	110	1	TRNK_MOUSE
28	6	5.0	113	1	TRNK_MOUSE
29	6	5.0	126	1	TRNK_MOUSE
30	6	5.0	130	1	TRNK_MOUSE
31	6	5.0	130	1	TRNK_MOUSE
32	6	5.0	141	1	TRNK_MOUSE
33	6	5.0	152	1	TRNK_MOUSE

34	6	5.0	153	1	MYG_HALGR	P02162 halichoerus
35	6	5.0	153	1	MYG_PROST	P30562 phoca sibir
36	6	5.0	153	1	MYG_PHYCA	P02185 physeter ca
37	6	5.0	154	1	RL2B_TOBAC	Q07761 nicotiana t
38	6	5.0	155	1	VG05_BPML5	Q05267 mycobacteri
39	6	5.0	156	1	CA20_HUMAN	P52298 homo sapien
40	6	5.0	160	1	Y650_TREPA	O83656 treponema p
41	6	5.0	162	1	YD73_METUA	O38768 methanococc
42	6	5.0	163	1	COAD_STRPY	P58104 streptococc
43	6	5.0	168	1	COX2_THERM	P98052 thermus aqu
44	6	5.0	173	1	TCPA_AMBME	P50157 ambystoma m
45	6	5.0	182	1	NUSG_CHLNR	O84322 chlamydia t

ALIGNMENTS

RESULT	1	TRNK_HUMAN	STANDARD:	PRT:	121 AA.
ID	TRNK_HUMAN				
AC	Q90HF0				
DR	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DE	16-OCT-2001 (Rel. 40, Last annotation update)				
GN	Neurokinin B precursor (NKB) (Neuromedin K) (ZNEUROK1).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Sheppard P., Jelinek L., Whitmore T., Blumberg H., Lehner J.,				
RA	O'Hara P.,				
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Placenta;				
RX	MEDLINE=20322570; PubMed=10866201;				
RA	Page N.M., Woods R.J., Gardiner S.M., Lomthasong K., Gladwell R.T.,				
RA	Butlin D.J., Manyonda I.T., Lowry P.J.;				
RT	"Excessive placental neurokinin B secretion during the third trimester				
RT	causes pre-eclampsia.";				
RL	Nature 405:797-800(2000).				
CC	- FUNCTION: TACHIKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,				
CC	EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND				
CC	SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH				
CC	MUSCLES (BY SIMILARITY).				
CC	- DEVELOPMENTAL STAGE: In pregnancy, the expression of NKB is				
CC	confined to the outer syncytiotrophoblast of the placenta,				
CC	significant concentrations of NKB can be detected in plasma as				
CC	early as week 9, and plasma concentrations of NKB are grossly				
CC	elevated in pregnancy-induced hypertension and pre-eclampsia.				
CC	- SIMILARITY: BELONGS TO THE TACHIKININ FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; AF161112; AA01430.1; -				
DR	EMBL; AF216586; AAF6980.1; -				
DR	EMBL; 162330; -				
DR	InterPro: IPR003635; Neurokinin.				
DR	InterPro: IPR002040; Tachikinin.				
DR	ProDom: PD020370; Neurokinin; 1.				
DR	ProSITE: PS00267; TACHIKININ; 1.				
KW	Tachikinin; Neuropeptide; Cleavage on pair of basic residues;				
KW	Amidation; Signal.				

```

FT SIGNAL 1 16 POTENTIAL.
FT PROPER 17 78 BY SIMILARITY.
FT PEPTIDE 81 90 NEUROKININ B.
FT PROPER 94 121 BY SIMILARITY.
FT MOD_RES 90 90 AMIDATION (G-91 PROVIDE AMIDE GROUP) (BY
SQ SEQUENCE 121 AA: 13438 MW: 14C9AFE2EE9EDECA CRC64;
SIMILARITY).

Query Match
Best Local Similarity 100.0%; Score 121; DB 1; Length 121;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRIMLFTAILAFSLAQSGAVCKEPEQEEVPGGSKRDPDLYOLLRLFKSHSLEGL 60
DB 1 MRIMLFTAILAFSLAQSGAVCKEPEQEEVPGGSKRDPDLYOLLRLFKSHSLEGL 60
OY 61 LKALSGASTDPKESTSPKRDMDHDFVGLMGRSVOPDSPFDVNOENVPSFGLTKYPPRA 120
DB 61 LKALSGASTDPKESTSPKRDMDHDFVGLMGRSVOPDSPFDVNOENVPSFGLTKYPPRA 120
OY 121 E 121
DB 121 E 121

RESULT 2
TKRK_MOUSE STANDARD: PRT: 116 AA.
AC P55099;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurokinin B precursor (NKB) (Neuromedin K) (Preprotachykinin B) (PPT-
B).
GN TAC3 OR NKMB OR TAC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR, TISSUE=Brain;
RA Kako K., Munekata E., Hosaka M., Murakami K., Nakayama K.;
RT "Cloning and sequence analysis of mouse CDMS encoding
preprotachykinin A and B."
RL Biomed. Res. 14:253-259(1993).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY). MANY SMOOTH
MUSCLES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D14423; BA003316.1; -.
CC MGD: MGI:98476; Tac2.
CC InterPro: IPR003635; Neurokinin.
CC InterPro: IPR002040; Tachykinin.
CC ProDom: PD020370; Neurokinin.
CC PROSITE: PS00267; TACHYKININ; 1.
CC Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
CC Amidation; signal.
FT SIGNAL 1 20 POTENTIAL.
FT PROPER 21 79
FT PEPTIDE 82 91 NEUROKININ B.
FT MOD_RES 91 91
SQ SEQUENCE 116 AA: 12659 MW: C73EC67F2BAF8C8C CRC64;

```

```

FT PROPER 95 116
FT MOD_RES 91 91 AMIDATION (G-92 PROVIDE AMIDE GROUP).
SQ SEQUENCE 116 AA: 12809 MW: BFE89373E2031CC CRC64;

Query Match
Best Local Similarity 100.0%; Score 15; DB 1; Length 116;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 KRDMHDFVGLMGR 93
DB 80 KRDMHDFVGLMGR 94

RESULT 3
TKRK_RAT STANDARD: PRT: 116 AA.
AC P08435;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurokinin B precursor (NKB) (Neuromedin K).
GN TAC3 OR NKMB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=88051833; PubMed=3479225;
RA Bonner T.I., Aftolter H.-U., Young A.C., Young W.S. III;
RT "A cDNA encoding the precursor of the rat neuropeptide, neurokinin
B."
RL Brain Res. 388:243-249(1987).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY). MANY SMOOTH
MUSCLES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M16410; AAA41711.1; -.
CC PIR: A43779; A43779.
CC InterPro: IPR003635; Neurokinin.
CC InterPro: IPR002040; Tachykinin.
CC ProDom: PD020370; Neurokinin.
CC PROSITE: PS00267; TACHYKININ; 1.
CC Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
CC Amidation; signal.
FT SIGNAL 1 20 POTENTIAL.
FT PROPER 21 79
FT PEPTIDE 82 91 NEUROKININ B.
FT PEPTIDE 95 116
FT MOD_RES 91 91
SQ SEQUENCE 116 AA: 12659 MW: C73EC67F2BAF8C8C CRC64;

Query Match
Best Local Similarity 100.0%; Score 15; DB 1; Length 116;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 KRDMHDFVGLMGR 93
DB 80 KRDMHDFVGLMGR 94

```



```

RESULT 4
ID TRNK_BOVIN STANDARD: PRT: 126 AA.
AC P08858:
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurokinin B precursor (NKB) (Neuromedin K).
GN TAC3 OR NKNB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86313713; PubMed=3462746;
RA Kotani H., Hoshimaru M., Nawa H., Nakanishi S.;
RT "Structure and gene organization of bovine neuromedin K precursor.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:7074-7078(1986).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: M14351; AAA30723.1; -
DR EMBL: M14347; AAA30723.1; JOINED.
DR EMBL: M14348; AAA30723.1; JOINED.
DR EMBL: M14349; AAA30723.1; JOINED.
DR EMBL: M14350; AAA30723.1; JOINED.
DR PIR: A25905; A25905.
DR InterPro: IPR003635; Neurokinin.
DR InterPro: IPR002040; Tachykinin.
DR ProDom: PD020370; Neurokinin.
DR PROSITE: PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
KW Amidation; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT PROPEP 21 83
FT PROPEP 86 95 NEUROKININ B.
FT PROPEP 99 126
FT MOD_RES 95 95 AMIDATION (G-96 PROVIDE AMIDE GROUP).
SQ SEQUENCE 126 AA; 13871 MW; 446EF43498EC059 CRC64;

Query Match 12.4%; Score 15; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 KRMHDFVGLMKR 93
DB 84 KRMHDFVGLMKR 98

RESULT 5
ID TRNK_PIG STANDARD: PRT: 10 AA.
AC P01292;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurokinin B (NKB) (Neuromedin K).

```

```

GN TAC3 OR NKNB.
OS Sus scrofa (Pig), and
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823, 8406;
RN [1]
RP SEQUENCE.
RX SPECIES-Pig; TISSUE-Spinal cord;
RX MEDLINE=83282812; PubMed=6576785;
RA Kangawa K., Minamino N., Fukuda A., Matsuo H.;
RT "Neuromedin K: a novel mammalian tachykinin identified in porcine
RT spinal cord.";
RL Biochem. Biophys. Res. Commun. 114:533-540(1983).
RN [2]
RP SEQUENCE.
RX SPECIES-R. ridibunda; TISSUE-Brain;
RX MEDLINE=92044543; PubMed=1658233;
RA O'Harte F., Burcher E., Lovaas S., Smith D.D., Vaudry H., Conlon J.M.;
RT "Ramkinin: a novel NK1 tachykinin receptor agonist isolated with
RT neurokinin B from the brain of the frog Rana ridibunda.";
RL J. Neurochem. 57:2086-2091(1991).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC CC
CC PIR: A01560; SPENK.
DR InterPro: IPR002040; Tachykinin.
DR PROSITE: PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1211 MW; E1FA62C9C9CA1 CRC64;

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. NO. 0.0004;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 DMHDFVGLM 90
DB 1 DMHDFVGLM 10

RESULT 6
ID SRPH_SYN7 STANDARD: PRT: 319 AA.
AC Q59967;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Serine acetyltransferase, plasmid (EC 2.3.1.30) (SAT).
GN SRPH.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95327059; PubMed=7603442;
RA Nicholson M.L., Gaasenbeek M., Laudenbach D.E.;
RT "Two enzymes together capable of cysteine biosynthesis are encoded on
RT a cyanobacterial plasmid.";
RL Mol. Gen. Genet. 247:623-632(1995).
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + L-serine -> CoA + O-acetyl-L-
CC serine.
CC -1- PATHWAY: CYSTEINE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NDL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL: U23436; AAA86726.1; -
 DR InterPro: IPR001451; Hexapep_transf.
 DR Pfam: PF00132; hexapep_3
 DR PROSITE: PS00101; HEXAPEP_TRANSFERRASES; FALSE_NEG.
 DR TRANSFERASE: Acyltransferase; Cysteine biosynthesis; Repeat; Plasmid.
 KW SEQUENCE 319 AA; 34570 MW; 00B874DBAAB1E70 CRC64;
 SQ

Query Match
 Best Local Similarity 7.4%; Score 9; DB 1; Length 319;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 ALSQASTDP 71
 Db 24 ALSQASTDP 32

 RESULT 7
 ID TKNL_RABIT STANDARD; PRT; 115 AA.
 AC P41540;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Protachykinin 1 precursor (PPT) [Contains: Substance P; Neurokinin A
 (NKA) (Substance K) (Neuromedin L); Neuropeptide gamma; C-terminal
 flanking peptide].
 GN TAC1 OR NKNA OR TAC2 OR NKA.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93371392; PubMed=8363593;
 RA Megeert H.J., Heitland A., Rose M., Forssmann W.G.;
 RT "Nucleotide sequence of the rabbit gamma-preprotachykinin I cDNA";
 RL Biochem. Biophys. Res. Commun. 195;128-131(1993).
 RN [2]
 RP SEQUENCE OF 72-92.
 RA Kage R., McGregor G.P., Thim L., Conlon J.M.;
 RT "Gamma-neuropeptide K: a peptide isolated from rabbit gut that is
 derived from gamma-preprotachykinin";
 RL Regul. Pept. 18;346-346(1987).
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS: ALPHA, BETA (SHOWN HERE),
 CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: X62964; CAAA4728.1; -
 DR PIR: S18922; S18922.
 DR InterPro: IPR003580; Protachykinin.
 DR InterPro: IPR002040; Tachykinin.
 DR Pfam: PF02202; Tachykinin; 1.
 DR

ProDom; PD005598; Protachykinin; 1.
 DR SMART: SM00203; TK; 2.
 DR PROSITE: PS00267; TACHYKININ; 2.
 KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
 KW Amidation; Alternative splicing; Signal; Neotransmitter.
 FT SIGMIL 1 19
 FT PEPTIDE 20 56
 FT PEPTIDE 58 68
 FT PEPTIDE 72 92
 FT PEPTIDE 83 92
 FT PEPTIDE 96 111
 FT MOD_RES 68 68
 FT MOD_RES 92 92
 SQ SEQUENCE 115 AA; 13370 MW; 5EC76F7C9B10E1C6 CRC64;
 SEC76F7C9B10E1C6 CRC64;

Query Match
 Best Local Similarity 6.6%; Score 8; DB 1; Length 115;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93
 Db 88 FVGLMGKR 95

 RESULT 8
 ID TKNL_HUMAN STANDARD; PRT; 129 AA.
 AC P20366; Q00072; O60600; O60601;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 16-FEB-1991 (Rel. 17, Last sequence update)
 DE Protachykinin 1 precursor (PPT) [Contains: Substance P; Neurokinin A
 (NKA) (Substance K) (Neuromedin L); Neuropeptide K (NPK); Neuropeptide
 gamma; C-terminal flanking peptide].
 GN TAC1 OR NKNA OR TAC2 OR NKA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM BETA).
 RX MEDLINE=87030957; PubMed=3770210;
 RA Harnar A.J., Armstrong A., Pascall J.C., Chapman K., Rosie R.,
 RA Curtis A., Going J., Edwards C.R.W., Fink G.;
 RT "cDNA sequence of human beta-preprotachykinin, the common precursor
 RT to substance P and neurokinin A";
 RL FEBS Lett. 208;67-72(1986).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM BETA).
 RC TISSUE=Brain;
 RA Tan A., Too H.P.;
 RT Submitted (Oct-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 36-122 FROM N.A. (ISOFORMS BETA AND GAMMA).
 RC TISSUE=Testis;
 RX MEDLINE=91209287; PubMed=1708336;
 RA Chivakata C., Brackmann B., Hunt N., Davidoff M., Schulze W.,
 RA Ivell R.;
 RT "Tachykinin (substance-P) gene expression in Leydig cells of the
 RT human and mouse testis";
 RL Endocrinology 128;2441-2448(1991).
 RN [4]
 RP SEQUENCE OF 98-107.
 RX MEDLINE=87275962; PubMed=3038549;
 RA Theodorsson-Norheim E., Joernvall H., Andersson M., Norheim I.,
 RA Oberg K., Jacobsson G.;
 RT "Isolation and characterization of neurokinin A, neurokinin A(3-10)
 RT and neurokinin A(4-10) from a neutral water extract of a metastatic
 RT ileal carcinoid tumour";
 RL Eur. J. Biochem. 166;693-697(1987).
 RN [5]
 RP SEQUENCE OF 36-118 FROM N.A. (ISOFORM ALPHA).
 RN

RC TISSUE=Blood, and Brain;
 RA Lai J.P., Douglas S.D., Rappaport E., Wu J.M., Ho W.Z.;
 RT "Identification of a delta isoform of preprotachykinin mRNA in human
 RL mononuclear phagocytes and lymphocytes."
 RN Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 [6]
 RP SEQUENCE OF 111-126.
 RC TISSUE=Adrenal medulla;
 RX MEDLINE-9113394; PubMed-2284201;
 RA Mcgregor G.P., Conlon J.M.;
 RT "Characterization of the C-terminal flanking peptide of human
 RL beta-preprotachykinin".
 CC Peptides 11:907-910(1990).
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS, ALPHA, BETA (SHOWN HERE),
 CC GAMMA AND DELTA, ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X54469; CAA38351.1; -
 DR EMBL; U37529; AAA79195.1; -
 DR EMBL; M68906; AAA60159.1; -
 DR EMBL; M68907; AAA60160.1; -
 DR EMBL; AF050656; AAC15702.1; -
 DR EMBL; AF050658; AAC15704.1; -
 DR PIR; A24805; A24805.
 DR PIR; S00069; S00069.
 DR MIM; 162320; -
 DR InterPro; IPR003580; Protachykinin.
 DR InterPro; IPR002040; Tachykinin.
 DR Pfam; PF02202; Tachykinin; 1.
 DR PRODOM; PD005598; Protachykinin; 1.
 DR SMART; SMO0203; TK; 2.
 DR PROSITE; PS00267; TACHYKININ; 2.
 KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
 KW Amdaation; Alternative splicing; Signal; Neurotransmitter.
 FT SIGNAL 1 19
 FT PROPEP 20 56
 FT PEPTIDE 58 68
 FT PEPTIDE 72 107
 FT PEPTIDE 72 107
 FT PEPTIDE 89 107
 FT PEPTIDE 98 107
 FT PEPTIDE 111 126
 FT MOD_RES 68 68
 FT MOD_RES 107 107
 FT VASAPLIC 74 88
 FT VASAPLIC 97 114
 FT VASAPLIC 115 115
 FT VASAPLIC 115 115
 FT CONFLICT 87 87
 FT SEQUENCE 129 AA; 15003 MW; 51412C1692368DE4 CRC64;
 Query Match 6.6%; Score 8; DB 1; Length 129;
 Best Local Similarity 100.0%; Pred. No. 0.47;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 86 FVGLMGR 93
 |||||
 Db 103 FVGLMGR 110

RESULT 9
 ID TKNL_BOVIN STANDARD; PRT: 130 AA.
 AC P01289; P01291; P04091; P20773;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protachykinin 1 precursor (PPT) [Contains: Substance P; Neurokinin A
 DE (NKA) (Substance K) (Neuromedin L); Neuropeptide K (NPK); Neuropeptide
 DE gamma; C-terminal flanking peptide];
 GN TAC1 OR NKNA OR TAC2 OR NKA.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM BETA).
 RX MEDLINE=85086245; PubMed=6083453;
 RA Nawa H., Kotani H., Nakanishi S.;
 RT "Tissue-specific generation of two preprotachykinin mRNAs from one
 RL gene by alternative RNA splicing".
 RN Nature 312:729-734(1984).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
 RX MEDLINE=84039802; PubMed=6195531;
 RA Nawa H., Hirose T., Takashima H., Inayama S., Nakanishi S.;
 RT "Nucleotide sequences of cloned cDNAs for two types of bovine brain
 RL substance P precursor".
 RN Nature 306:32-36(1983).
 [3]
 RP SEQUENCE OF 36-122 FROM N.A. (ISOFORMS BETA AND GAMMA).
 RC TISSUE=Hypothalamus;
 RX MEDLINE=91209287; PubMed=1708336;
 RA Chivakata C., Brackmann B., Hunt N., Davidoff M., Schulze W.,
 RA Iveli R.;
 RT "Tachykinin (substance-P) gene expression in Leydig cells of the
 RL human and mouse testis".
 RL Endocrinology 128:2441-2448(1991).
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS, ALPHA, BETA (SHOWN HERE),
 CC GAMMA AND DELTA, ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X00075; CAA24939.1; -
 DR EMBL; X00075; CAA24940.1; -
 DR EMBL; X00075; CAA24941.1; -
 DR EMBL; X00076; CAA24942.1; -
 DR EMBL; X00076; CAA24943.1; -
 DR EMBL; X02351; CAA26206.1; ALT_SEQ.
 DR EMBL; X02351; CAA26206.1; JOINED.
 DR EMBL; X01396; CAA26206.1; JOINED.
 DR EMBL; X01396; CAA26206.1; JOINED.
 DR EMBL; X01398; CAA26206.1; JOINED.
 DR EMBL; X01399; CAA26206.1; JOINED.
 DR EMBL; X01400; CAA26206.1; JOINED.
 DR EMBL; M68911; AAA30724.1; -
 DR EMBL; M68912; AAA30725.1; -
 DR PIR; A01557; SPBOA.
 DR PIR; A01559; SPBOB.
 DR PIR; A05093; A05093.

DR PIR; B25067; B25067.
 DR InterPro: IPR003580; Protachykinin.
 DR InterPro: IPR002040; Tachykinin.
 DR Pfam: PF02202; Tachykinin; 1.
 DR PRODOM; PD005598; Protachykinin; 1.
 DR SMART; SM00203; TK; 2.
 DR PROSITE; PS00267; TACHYKININ; 2.
 DR Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
 KW Amidation; Alternative splicing; Signal; Neurotransmitter.
 FT SIGNAL 1 19
 FT PROPEP 20 56
 FT PEPTIDE 58 107
 FT PEPTIDE 72 107
 FT PEPTIDE 72 107
 FT PEPTIDE 72 107
 FT PEPTIDE 89 107
 FT PEPTIDE 98 107
 FT PEPTIDE 107 126
 FT PEPTIDE 111 126
 FT MOD.RES 68 68
 FT MOD.RES 107 107
 FT VARSPLIC 74 88
 FT VARSPLIC 97 114
 FT VARSPLIC 115 115
 FT CONFLICT 121 121
 FT SEQUENCE 130 AA; 15076 MW; CE2A28572305DEB7 CRC64;

Query Match
 Best Local Similarity 6.6%; Score 8; DB 1; Length 130;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 86 FVGLMGKR 93
 |||||
 DB 103 FVGLMGKR 110

RESULT 10
 ID TKNL_MESAU STANDARD; PRT; 130 AA.
 AC 060541; P49110;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protachykinin 1 precursor (PP1) [Contains: Substance P; Neurokinin A
 DE (NKA) (Substance K) (Neuromedin L); Neuropeptide K (NPK); Neuropeptide
 DE gamma; C-terminal flanking peptide].
 GN TAC1 OR NKMA OR TAC2 OR NKA.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS BETA AND GAMMA).
 RC STRAIN=AURA; TISSUE=Brain;
 RA Heiland A., Kruboffer M., Juergen Maegert H.J., Forssmann W.G.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),
 CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC EMBL; X80662; CAAS6691.1; -
 CC EMBL; X80663; CAAS6692.1; -
 DR InterPro: IPR003580; Protachykinin.
 DR InterPro: IPR002040; Tachykinin.
 DR Pfam: PF02202; Tachykinin; 1.
 DR PRODOM; PD005598; Protachykinin; 1.
 DR SMART; SM00203; TK; 2.
 DR PROSITE; PS00267; TACHYKININ; 2.
 DR Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
 KW Amidation; Alternative splicing; Signal; Neurotransmitter.
 FT SIGNAL 1 19
 FT PROPEP 20 56
 FT PEPTIDE 58 107
 FT PEPTIDE 72 107
 FT PEPTIDE 72 107
 FT PEPTIDE 72 107
 FT PEPTIDE 89 107
 FT PEPTIDE 98 107
 FT PEPTIDE 111 126
 FT MOD.RES 68 68
 FT MOD.RES 107 107
 FT VARSPLIC 74 88
 FT VARSPLIC 97 114
 FT SEQUENCE 130 AA; 14907 MW; CC92E9371A646F2E CRC64;

Query Match
 Best Local Similarity 6.6%; Score 8; DB 1; Length 130;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 86 FVGLMGKR 93
 |||||
 DB 103 FVGLMGKR 110

RESULT 11
 ID TKNL_MOUSE STANDARD; PRT; 130 AA.
 AC P41539; Q00073;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protachykinin 1 precursor (PP1) [Contains: Substance P; Neurokinin A
 DE (NKA) (Substance K) (Neuromedin L); Neuropeptide K (NPK); Neuropeptide
 DE gamma; C-terminal flanking peptide].
 GN TAC1 OR NKMA OR TAC2 OR NKA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM BETA).
 RC STRAIN=ICR; TISSUE=Brain;
 RA Kato K., Muneakata E., Hosaka M., Murakami K., Nakayama K.;
 RT "Cloning and sequence analysis of mouse cDNAs encoding
 RT preprotachykinin A and B.";
 RL Biomed. Res. 14:253-259(1993).
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),
 CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

```

-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL: D17584; BAA04508.1; -
DR      EMBL: M68908; AAA39869.1; -
DR      EMBL: M68909; AAA39870.1; -
DR      MGD: MGI:98474; Tacl.
DR      InterPro: IPR003580; Protachykinin.
DR      InterPro: IPR002040; Tachykinin.
DR      Pfam: PF02202; Tachykinin.1.
DR      ProDom: PD005598; Protachykinin.1.
DR      SMART: SM00203; TK; 2.
DR      PROSITE: PS00267; TACHYKININ; 2.
DR      Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
DR      Amidation; Alternative splicing; Signal; Neurotransmitter.
KW      SIGNAL.
FT      PROPEP      20      56      SUBSTANCE P.
FT      PEPTIDE      58      68      NEUROPEPTIDE K.
FT      PEPTIDE      72      107      NEUROPEPTIDE GAMMA 1ST PART.
FT      PEPTIDE      72      73      NEUROPEPTIDE GAMMA 2ND PART.
FT      PEPTIDE      89      107      NEUROPEPTIDE GAMMA 2ND PART.
FT      PEPTIDE      98      107      NEUROKININ A.
FT      PEPTIDE      111      126      C-TERMINAL FLANKING PEPTIDE (POTENTIAL).
FT      MOD_RES      68      68      AMIDATION (G-69 PROVIDE AMIDE GROUP).
FT      MOD_RES      107      107      AMIDATION (G-108 PROVIDE AMIDE GROUP).
FT      VASPLIC      74      88      MISSING (IN ISOFORM GAMMA).
SO      SEQUENCE      130 AA; 15045 MW; 7BE8DA15FDE72FF8 CRC64;

Query Match      6.6%; Score 8; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      86      FVGLMGKR 93
DB      103      FVGLMGKR 110

RESULT 12
TKN1_RAT
AC      P06767; P08856; P08857; P22356; PRT; 130 AA.
DT      01-JAN-1988 (Rel. 06; Created)
DT      01-NOV-1988 (Rel. 09; Last sequence update)
DT      16-OCT-2001 (Rel. 40; Last annotation update)
DE      Protachykinin 1 precursor (PPT) [contains: Substance P; Neurokinin A
DE      (NKA) (Substance K) (Neurokinin B); Neuropeptide K (NPK); Neuropeptide
DE      gamma; C-terminal flanking peptide].
GN      TAC1 OR NKNA OR TAC2 OR NKA.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_Taxid=10116;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA AND GAMMA).
RX      MEDLINE-90331040; PubMed-165945;
RA      Carter M.S., Krause J.E.;
RT      "Structure, expression, and some regulatory mechanisms of the rat
RT      preprotachykinin gene encoding substance P, neurokinin A,
RT      neuropeptide K, and neuropeptide gamma.";
RL      J. Neurosci. 10:2203-2214(1990).
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA AND GAMMA).
RX      MEDLINE-87118268; PubMed-2433692;
RA      Krause J.E., Chigwin J.M., Carter M.S., Xu Z.S., Hershey A.D.;
RT      "Three rat preprotachykinin mRNAs encode the neuropeptides substance
RT      P and neurokinin A.";
```

```

RL      Proc. Natl. Acad. Sci. U.S.A. 84:881-885(1987).
RN      [3]
RP      SEQUENCE FROM N.A. (ISOFORM GAMMA).
RX      MEDLINE-87025808; PubMed-2429656;
RA      Kawaguchi Y., Hoshimaru M., Nawa H., Nakanishi S.;
RT      "Sequence analysis of cloned cDNA for rat substance P precursor:
RT      existence of a third substance P precursor.";
RL      Biochem. Biophys. Res. Commun. 139:1040-1046(1986).
RN      [4]
RP      SEQUENCE FROM N.A. (ISOFORM DELTA).
RC      TISSUE-Dorsal root ganglion;
RX      MEDLINE-91085565; PubMed-17702066;
RA      Harmar A.J., Hyde V., Chapman K.E.;
RT      "Identification and cDNA sequence of delta-preprotachykinin, a fourth
RT      splicing variant of the rat substance P precursor.";
RL      FEBS Lett. 275:22-24(1990).
RN      [5]
RP      SEQUENCE OF 1-41 FROM N.A.
RX      MEDLINE-93192337; PubMed-8448217;
RA      Chapman K.E., Lyons V., Harmar A.J.;
RT      "The sequence of 5' flanking DNA from the rat preprotachykinin gene;
RT      analysis of putative transcription factor binding sites.";
RL      Biochim. Biophys. Acta 1172:361-363(1993).
CC      -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC      EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC      SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC      MUSCLES.
CC      -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS: ALPHA, BETA (SHOWN HERE),
CC      GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC      -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL: M34162; AAA41926.1; -
DR      EMBL: M34159; AAA41926.1; JOINED.
DR      EMBL: M34160; AAA41926.1; JOINED.
DR      EMBL: M34161; AAA41926.1; JOINED.
DR      EMBL: M34184; AAA41925.1; -
DR      EMBL: M34183; AAA41929.1; -
DR      EMBL: M15191; AAA41928.1; -
DR      EMBL: M14312; AAA41927.1; -
DR      EMBL: L07328; AAA41924.1; -
DR      EMBL: X56306; CAA39752.1; -
DR      PIR: A26590; A26590.
DR      PIR: B26590; B26590.
DR      PIR: C26590; C26590.
DR      PIR: A37163; A37163.
DR      PIR: S12958; S12958.
DR      InterPro: IPR003580; Protachykinin.
DR      InterPro: IPR002040; Tachykinin.
DR      Pfam: PF02202; Tachykinin.1.
DR      PROSITE: PD005598; Protachykinin.1.
DR      SMART: SM00203; TK; 2.
DR      PROSITE: PS00267; TACHYKININ; 2.
DR      Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
DR      Amidation; Alternative splicing; Signal; Neurotransmitter.
KW      SIGNAL.
FT      PROPEP      1      19      POTENTIAL.
FT      PEPTIDE      20      56      SUBSTANCE P.
FT      PEPTIDE      58      68      NEUROPEPTIDE K.
FT      PEPTIDE      72      107      NEUROPEPTIDE GAMMA 1ST PART.
FT      PEPTIDE      72      73      NEUROPEPTIDE GAMMA 2ND PART.
FT      PEPTIDE      89      107      NEUROKININ A.
FT      PEPTIDE      98      107      C-TERMINAL FLANKING PEPTIDE (POTENTIAL).
FT      MOD_RES      111      126      AMIDATION (G-69 PROVIDE AMIDE GROUP).
FT      MOD_RES      68      68      AMIDATION (G-108 PROVIDE AMIDE GROUP).
FT      VASPLIC      107      107      MISSING (IN ISOFORM GAMMA AND ISOFORM
FT      VASPLIC      74      88
```

FT VARSPLIC 97 114 DELTA) :
 FT MISSING (IN ISOFORM ALPHA AND ISOFORM
 FT VARSPLIC 115 115 DELTA).
 FT V -> M (IN ISOFORM ALPHA AND ISOFORM
 SO SEQUENCE 130 AA; 15001 MM; B2EFEB60DCD75A CRC64;

Query Match 6.6%; Score 8; DB 1; Length 130;
 Best Local Similarity 100.0%; Pred. No. 0.47;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93
 |||||
 Db 103 FVGLMGKR 110

RESULT 13
 ID POLG_PEMVM STANDARD; PRT: 3099 AA.
 AC 056075;

DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polypeptide [contains: N-terminal protein (P1); Helper
 DE component polypeptide (EC 3.4.22.-) (HC-PRO); Protein P3; 6 kDa protein
 DE 1 (6K1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2 (6K2);
 DE Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A) (NI)
 DE (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-PRO); Nuclear inclusion
 DE protein B (NI-B) (NI) (RNA-directed RNA polymerase) (EC 2.7.7.48);
 DE Coat protein (CP);
 OS Peanut mottle virus (strain M).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
 OC Polyviruses.
 OX NCBI_TaxID-103926;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Flasiński S., Gonzales R.A., Cassidy B.G.;
 RT "The complete nucleotide sequence of peanut mottle virus (M strain)
 RL genomic RNA.";
 CC Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
 CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
 CC MAY BE INVOLVED IN REPLICATION.
 CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: Hydrolyses glutamyl bonds, and activity is
 CC further restricted by preferences for the amino acids in P6 - P1,
 CC that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
 CC Glu(Ser or Gly) for the enzyme from tobacco etch virus. The
 CC natural substrate is the viral polypeptide, but other proteins and
 CC oligopeptides containing the appropriate consensus sequence are
 CC also cleaved.
 CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
 CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
 CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF023848; AAB94595.1; -;
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.

DR InterPro: IPR001730; Peptidase_C4.
 DR InterPro: IPR001456; Peptidase_C6.
 DR InterPro: IPR002540; Poty_P1.
 DR InterPro: IPR001592; Poty_P1.
 DR InterPro: IPR001205; RNA_Pol_P3D.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR Pfam: PF00863; Peptidase_C4; 1.
 DR Pfam: PF00851; Peptidase_C6; 1.
 DR Pfam: PF00767; Poty-coat; 1.
 DR Pfam: PF01577; Poty_P1; 1.
 DR Pfam: PF06880; RNA_dep_RNA_pol; 1.
 DR PRINTS: PR00956; NIAPOTYPTASE.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR Hydrolase: Transferase; Thiol protease; RNA-directed RNA polymerase;
 KW Coat protein; Polypeptide; Covalent protein-RNA linkage; Helicase;
 KW ATP-binding.
 KM CHAIN 1 322 N-TERMINAL PROTEIN.
 FT CHAIN 323 779 HELPER COMPONENT PROTEINASE.
 FT CHAIN 780 1128 PROTEIN P3.
 FT CHAIN 1129 1180 6 KDA PROTEIN 1.
 FT CHAIN 1181 1814 CYTOPLASMIC INCLUSION PROTEIN.
 FT CHAIN 1815 1867 6 KDA PROTEIN 2.
 FT CHAIN 1868 ? GENOME-LINKED PROTEIN.
 FT CHAIN ? 2303 NUCLEAR INCLUSION PROTEIN A.
 FT CHAIN 2304 2821 NUCLEAR INCLUSION PROTEIN B.
 FT CHAIN 2822 3099 COAT PROTEIN.
 FT CHAIN 2828 2303 PUTATIVE NUCLEAR INCLUSION PROTEIN A.
 FT NP_BIND 1265 1272 ATP (POTENTIAL).
 SO SEQUENCE 3099 AA; 351032 MM; 0DBE9FC7603F04B CRC64;

Query Match 6.6%; Score 8; DB 1; Length 3099;
 Best Local Similarity 100.0%; Pred. No. 8.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

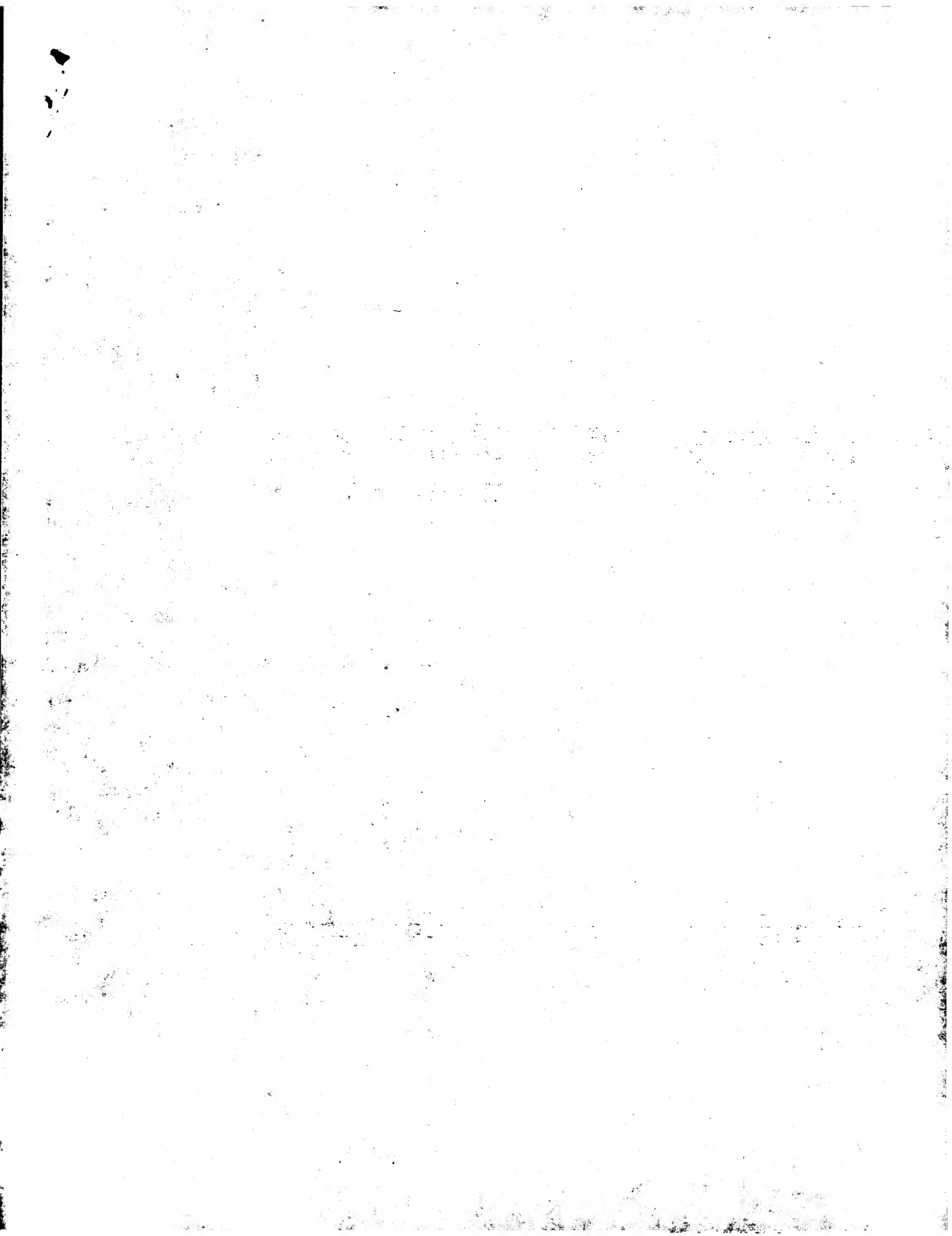
QY 28 EEVPGGG 35
 |||||
 Db 546 EEVPGGG 553

RESULT 14
 ID YN00_MYCTU STANDARD; PRT: 310 AA.
 AC 050665;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Hypothetical 34.4 kDa protein Rv2300c.
 DE Rv2300c OR M72357 OR M7CY339.09.
 OS Mycobacterium tuberculosis.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID-1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE-98295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., Mclean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Sultun J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z77163; CAB0971.1; -;
 DR EMBL; AE007078; AAK4642.1; -;
 DR TIGR; MT2357; -;
 DR Tuberculist; RV2300C; -;
 DR InterPro: IPR001279; Beta_lactam_met.
 DR Pfam; PF00753; lactamase_B; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 310 AA: 34352 MW: 0764F4F64AB5E51 CRC64;
 OY 40 DEDLYOL 46
 DB 239 DEDLYOL 305
 Query Match 5.8%; Score 7; DB 1; Length 310;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 15
 GRS5_BOVIN STANDARD; PRT; 590 AA.
 ID GRS5_BOVIN
 AC P43249;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE G protein-coupled receptor kinase GRS5 (EC 2.7.1.-).
 GN GPRK5 OR GRS5.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE= Tongue epithelium;
 RX MEDLINE=94165084; PubMed=8120045;
 RA Piemont R.T., Koch W.J., Inglese J., Lefkowitz R.J.;
 RT "Identification, purification, and characterization of GRS5, a member
 RT of the family of G protein-coupled receptor kinases.";
 RL J. Biol. Chem. 269:6832-6841(1994).
 CC -!- FUNCTION: SPECIFICALLY PHOSPHORYLATES THE ACTIVATED FORMS OF G
 CC PROTEIN-COUPLED RECEPTORS.
 CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN LUNG, HEART, RETINA, LINGUAL
 CC EPITHELIUM. VERY LITTLE IN BRAIN, LIVER, KIDNEY.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC GPRK SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U01206; AAA17561.1; -;
 DR HSSP; P05132; IBKX.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR000239; GPCR_kinase.
 DR InterPro: IPR000961; PKinase_C.
 DR InterPro: IPR000342; RGS.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; PKinase; 1.
 DR Pfam; PF00615; RGS; 1.
 DR PRINTS; PRO0717; GPCR_KINASE.
 DR SMART; SM00315; RGS; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS0132; RGS; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation.
 FT DOMAIN 1 185 N-TERMINAL.
 FT DOMAIN 186 448 PROTEIN_KINASE.
 FT DOMAIN 449 590 C-TERMINAL.
 FT DOMAIN 53 171 RGS.. (BY SIMILARITY).
 FT NP_BIND 192 200 ATP (BY SIMILARITY).
 FT BINDING 215 215 BY SIMILARITY.
 FT ACT_SITE 311 311 PHOSPHORYLATION (AUTO-).
 FT MOD_RES 484 484 PHOSPHORYLATION (AUTO-).
 FT MOD_RES 485 485
 SQ SEQUENCE 590 AA: 67688 MW: EBF353697D8BA6E2 CRC64;
 OY 46 LIQRLEK 52
 DB 550 LIQRLEK 556
 Query Match 5.8%; Score 7; DB 1; Length 590;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: May 24, 2002, 17:07:42
 Job time: 126 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 17:04:56 ; Search time 16.07 Seconds
(without alignments)
723.510 Million cell updates/sec

Title: US-09-852-659-68
121

Perfect score: 121
Sequence: 1 MRIMLFTAILAFSLAOSFG.....DVGQENVSPGILKYPPEAE 121

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	15	12.4	116	2 A43779	neurokinin B precu
2	15	12.4	116	2 I65342	tachykinin B precu
3	15	12.4	126	2 A25905	tachykinin B precu
4	10	8.3	10	1 SPECNK	neurokinin K - pig
5	9	7.4	319	2 S53322	srph protein
6	8	6.6	63	2 JC2412	tachykinin gamma c
7	8	6.6	72	2 I62742	tachykinin A gamma
8	8	6.6	72	2 JC5455	preprotachykinin-A
9	8	6.6	115	1 SPBRG	substance P gamma
10	8	6.6	115	2 S47039	tachykinin 1 precu
11	8	6.6	129	1 SPHOB	neurokinin 1 precu
12	8	6.6	130	1 SPRTB	substance P beta p
13	8	6.6	130	1 SPBOB	neurokinin 1 precu
14	8	6.6	130	1 S47038	tachykinin 1 precu
15	8	6.6	130	2 I52526	neurokinin 1 precu
16	8	6.6	812	2 E75338	DNA gyrase, subun
17	8	6.6	835	2 AH1861	DNA topoisomerase
18	7	5.8	143	2 S75133	hypothetical prote
19	7	5.8	252	2 AC2226	hypothetical prote
20	7	5.8	276	2 AD3581	transcription regu
21	7	5.8	310	2 H70733	hypothetical prote
22	7	5.8	366	2 T50468	probable maturase
23	7	5.8	377	2 B82846	succinyl-diaminop
24	7	5.8	394	2 E75439	conserved hypochet
25	7	5.8	475	2 C81351	probable UDP-N-ace
26	7	5.8	490	2 H95033	xanthine/uracil pe
27	7	5.8	490	2 H97904	conserved hypochet
28	7	5.8	493	2 E96664	hypothetical prote
29	7	5.8	496	2 C83122	probable aldehyde

30	7	5.8	534	2 D83775	hypothetical prote
31	7	5.8	590	1 A54372	G protein-coupled
32	7	5.8	590	2 A48277	G protein-coupled
33	7	5.8	619	2 B90141	hypothetical prote
34	7	5.8	637	2 B83052	Dnak protein PA476
35	7	5.8	648	1 A64167	cytochrome c-type
36	7	5.8	654	2 B84680	hypothetical prote
37	7	5.8	702	2 A86383	76.4K protein kina
38	7	5.8	712	2 T33231	hypothetical prote
39	7	5.8	824	2 F87330	photosynthesis pro
40	7	5.8	1038	2 T51518	ubiquitin-fusion d
41	7	5.8	1058	2 T50496	hypothetical prote
42	7	5.8	1516	2 F84919	glutathione-conjug
43	6	5.0	20	2 PC2084	serine proteinase
44	6	5.0	35	2 P50439	potassium channel
45	6	5.0	72	1 B24033	small acid-soluble

ALIGNMENTS

RESULT 1
A43779
neurokinin B precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 16-Jul-1999
C:Accession: A43779
R:Bonner, T.I.; Affolter, H.U.; Young, A.C.; Young III, W.S.
Brain Res. Mol. Brain Res. 2, 243-249, 1987
A:Title: A cDNA encoding the precursor of the rat neuropeptide, neurokinin B.
A:Reference number: A43779
A:Accession: A43779
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-116 <BON>
A:Cross-references: GB:M16410; GB:M35607; NID:g205724; PIDN:AAA41711.1; PID:g205725
C:Superfamily: neurokinin B precursor
C:Keywords: neuropeptide

Query Match 12.4%; Score 15; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 KRDMHDFVGLMGKR 93
DB 80 KRDMHDFVGLMGKR 94

RESULT 2
I65342
tachykinin B precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I65342
R:Nako, K.; Muneke, E.; Hosaka, M.; Murakami, K.; Nakayama, K.
Biomed. Res. 14, 251-259, 1993
A:Title: Cloning and sequence analysis of mouse cDNAs encoding preprotachykinin A and
A:Reference number: I52526
A:Accession: I65342
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-116 <RES>
A:Cross-references: GB:D14423; NID:g407346; PIDN:BAA03316.1; PID:g407347
C:Genetics:
A:Gene: PPT-B
C:Superfamily: neurokinin B precursor

Query Match 12.4%; Score 15; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 KRDHDFVGLMGKR 93
|||||
Db 80 KRDHDFVGLMGKR 94

RESULT 3
A25905
tachykinin B precursor - bovine
N:Alternate names: neuromedin K
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A25905
R:Kotani, H.; Hoshimaru, M.; Nawa, H.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 83, 7074-7078, 1986
A:Title: Structure and gene organization of bovine neuromedin K precursor.
A:Reference number: A25905; MUID:86313713
A:Accession: A25905
A:Molecule type: DNA
A:Residues: 1-126 <NOT>
A:Cross-references: GB:M14351; NID:g163587; PIDN:AAA30723.1; PID:g163590
C:Superfamily: neurokinin B precursor
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-126/Product: protachykinin B #status predicted <MAT>

Query Match 12.4%; Score 15; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 79 KRDHDFVGLMGKR 93
|||||
Db 84 KRDHDFVGLMGKR 98

RESULT 4
SPGCK
neuromedin K - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 23-Aug-1996
C:Accession: A01560
R:Kangawa, K.; Minamino, N.; Fukuda, A.; Matsuo, H.
Biochem. Biophys. Res. Commun. 114, 533-540, 1983
A:Title: Neuromedin K: a novel mammalian tachykinin identified in porcine spinal cord.
A:Reference number: A01560; MUID:83282812
A:Accession: A01560
A:Molecule type: protein
A:Residues: 1-10 <KRN>
A:Note: The structure of the peptide was confirmed by synthesis
C:Comment: The biological source of this peptide is spinal cord. It stimulates smooth m
C:Superfamily: neurokinin B precursor
C:Keywords: amidated carboxyl end; hormone; spinal cord
F:10/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 81 DMHDFVGLM 90
|||||
Db 1 DMHDFVGLM 10

RESULT 5
SS5322
srpH protein - *Synechococcus* sp. (strain PCC 7942) plasmid pANL
C:Species: *Synechococcus* sp.
A:Variety: PCC 7942
C:Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 18-Jun-1999
C:Accession: SS5322
R:Nicholson, M.L.; Gaasenbeek, M.; Laudenbach, D.E.
Mol. Gen. Genet. 247, 623-632, 1995
A:Title: Two enzymes together capable of cysteine biosynthesis are encoded on a cyanobac

A:Reference number: SS5321; MUID:95327059
A:Accession: SS5322
A:Molecule type: DNA
A:Residues: 1-319 <NIC>
C:Genetics:
A:Gene: srpH
A:Gene: plasmid pANL
C:Superfamily: serine acetyltransferase homology
F:142-311/Domain: serine acetyltransferase homology <SAR>

Query Match 7.4%; Score 9; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 63 ALSQASTDP 71
|||||
Db 24 ALSQASTDP 32

RESULT 6
JC2412
tachykinin gamma chain precursor - rat
C:Species: *Rattus norvegicus* (Norway rat)
C:Date: 25-Feb-1995 #sequence_revision 26-May-1995 #text_change 17-Mar-1999
C:Accession: JC2412
R:Khan, I.; Collins, S.M.
Biochem. Biophys. Res. Commun. 202, 796-802, 1994
A:Title: Fourth isoform of preprotachykinin messenger RNA encoding for substance P in
A:Reference number: JC2411; MUID:94324969
A:Accession: JC2412
A:Molecule type: mRNA
A:Residues: 1-63 <KHA>
C:Superfamily: substance P precursor
C:Keywords: amidated carboxyl end
F:12-21/Product: substance P #status predicted <SUP>
F:21/Modified site: amidated carboxyl end (Met) (amide in mature form from following

Query Match 6.6%; Score 8; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 86 FVGLMGKR 93
|||||
Db 41 FVGLMGKR 48

RESULT 7
162742
tachykinin A gamma chain precursor - mouse (fragment)
C:Species: *Mus musculus* (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: 162742; JC5453
R:Chiwakata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schulze, W.; Iwell, R.
Endocrinology 128, 2441-2448, 1991
A:Title: Tachykinin (substance-P) gene expression in Leydig cells of the human and mo
A:Reference number: JC5450; MUID:91209287
A:Accession: 162742
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-72 <RES>
A:Cross-references: GB:M68909; NID:g200469; PIDN:AAA39970.1; PID:g554261
C:Comment: This protein contains two tachykinin peptide hormone substance-P which is
C:Genetics:
A:Gene: gamma-PPT-A
C:Superfamily: substance P precursor
F:1-72/Domain: signal sequence #status predicted <SIG>
F:73-33/Product: substance-P #status predicted <SRP>
F:48-57/Product: neurokinin-A #status predicted <NKA>

Query Match 6.6%; Score 8; DB 2; Length 72;

Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 FVGLMGKR 93
|||||
DB 53 FVGLMGKR 60

RESULT 8

JC5455
preprotachykinin-A gamma precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 16-Jul-1999
C:Accession: JC5455; I45967
R:Chikata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schultze, W.; Ivell, R.
Endocrinology 128, 2441-2448, 1991
A:Title: Tachykinin (substance-P) gene expression in Leydig cells of the human and mouse
A:Reference number: JC5450; MUID:91209287
A:Accession: JC5455
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-72 <CHT>
A:Cross-references: GB:M68912; NID:g163593; PIDN:AA30725.1; PID:g552336
C:Comment: This protein contains two tachykinin peptide hormone substance-P which is inv
C:Genetics:
A:Gene: PPT-A
C:Superfamily: substance P precursor
F:1-72/Domain: signal sequence #status predicted <SIG>
F:23-33/Product: substance-P #status predicted <STP>
F:48-57/Product: neurokinin-A #status predicted <NKA>

Query Match 6.6%; Score 8; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93
|||||
DB 53 FVGLMGKR 60

RESULT 9

SPRBG
substance P gamma precursor - rabbit
N:Alternate names: gamma-neuropeptide K; gamma-preprotachykinin I precursor; tachykinin
N:Contains: neurokinin A; neuropeptide K; substance P
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 10-Nov-1992 #sequence_revision 26-May-1995 #text_change 18-Jun-1999
C:Accession: JN0709; A60302; A60200; S18922
R:Maegert, H.J.; Heitland, A.; Rose, M.; Forssmann, W.G.
Biochem. Biophys. Res. Commun. 195, 128-131, 1993
A:Title: Nucleotide sequence of the rabbit gamma-preprotachykinin I cDNA.
A:Reference number: JN0709; MUID:93371392
A:Accession: JN0709
A:Molecule type: mRNA
A:Residues: 1-115 <MA>
A:Cross-references: EMBL:X62994; NID:g1565; PIDN:CAA44728.1; PID:g1566
R:Kage, R.; McGregor, G.P.; Thim, L.; Conlon, J.M.
Regul. Pept. 18, 346, 1987
A:Title: gamma-neuropeptide K: a peptide isolated from rabbit gut that is derived from g
A:Reference number: A60302
A:Accession: A60302
A:Molecule type: protein
A:Residues: 72-92 <KAG>
J:Neurochem. 50, 1412-1417, 1988
A:Title: Neuropeptide-gamma: a peptide isolated from rabbit intestine that is derived fr
A:Reference number: A60200; MUID:88199570
A:Accession: A60200
A:Molecule type: protein
A:Residues: 72-92 <KA>
C:Comment: The gamma alternatively spliced form is processed to yield substance P and ne
C:Superfamily: substance P precursor

C:Keywords: alternative splicing; amidated carboxyl end; hormone; neuropeptide; tachy
F:1-15/Domain: signal sequence #status predicted <SIG>
F:58-68/Product: substance P #status predicted <SBP>
F:72-92/Product: gamma-neuropeptide K #status experimental <NPK>
F:83-92/Product: neurokinin A #status predicted <NKA>
F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following
F:92/Modified site: amidated carboxyl end (Met) (amide in mature form from following

Query Match 6.6%; Score 8; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93
|||||
DB 88 FVGLMGKR 95

RESULT 10

S47039
tachykinin 1 precursor - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C:Accession: S47039
R:Heitland, A.; Kruhoffer, M.; Juergen Maegert, H.J.; Forssmann, W.G.
submitted to the EMBL Data Library, July, 1994
A:Reference number: S47038
A:Accession: S47039
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-115 <HE>
A:Cross-references: EMBL:X80663; NID:g520938; PIDN:CAA56692.1; PID:g520939
C:Superfamily: substance P precursor

Query Match 6.6%; Score 8; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93
|||||
DB 88 FVGLMGKR 95

RESULT 11

SPHUB
neurokinin 1 precursor, beta splice form [validated] - human
N:Alternate names: neurokinin A; neurokinin alpha; neuromedin L; neuropeptide K; prep
N:Contains: neurokinin 1; neurokinin 1 precursor, alpha splice form; neurokinin 1 pre
C:Species: Homo sapiens (man)
C:Date: 12-Feb-1988 #sequence_revision 26-May-1995 #text_change 19-May-2000
C:Accession: A24805; A60425; S00065; S03033; JC5451; JC5450; A59270; B59270;
R:Hammar, A.J.; Armstrong, A.; Pascall, J.C.; Chapman, K.; Rosie, R.; Curtis, A.; Gol
FEBS Lett. 208, 67-72, 1986
A:Title: cDNA sequence of human beta-preprotachykinin, the common precursor to substa
A:Reference number: A24805; MUID:87030957
A:Accession: A24805
A:Molecule type: mRNA
A:Residues: 1-129 <HAB>
A:Cross-references: GB:M28109; EMBL:X54469; NID:g29482; PIDN:CAA38351.1; PID:g29483
R:McGregor, G.P.; Conlon, J.M.
Peptides 11, 907-910, 1990
A:Title: Characterization of the C-terminal flanking peptide of human beta-preprotach
A:Reference number: A60425; MUID:91133994
A:Accession: A60425
A:Molecule type: protein
A:Residues: 111-126 <MCG>
A:Experimental source: neuroendocrine tumor of adrenal medulla
R:Theodorsson-Norheim, E.; Joernvall, H.; Andersson, M.; Norheim, I.; Oberg, K.; Jac
Eur. J. Biochem. 166, 693-697, 1987
A:Title: Isolation and characterization of neurokinin A, neurokinin A(3-10) and neuro
A:Reference number: S00069; MUID:87275962
A:Accession: S00069

A:Molecule type: protein
 A:Residues: 98-107 <THE>
 R:Kage, R.; Thim, L.; Creutzfeldt, W.; Conlon, J.M.
 Biochem. J. 253, 203-207, 1988
 A:Title: Post-translational processing of preprotachykinins. Isolation of protachykinin-
 A:Reference number: 503033; MUID:8839887
 A:Accession: 503033
 A:Molecule type: protein
 A:Residues: 20-30 <KAG>
 R:Chiwakata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schulze, W.; Iwell, R.
 Endocrinology 128, 2441-2448, 1991
 A:Title: Tachykinin (substance-P) gene expression in Leydig cells of the human and mouse
 A:Reference number: JC3450; MUID:91209287
 A:Accession: JC3451
 A:Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 36-73,89-122 <CHI1>
 A:Cross-references: GB:M68907; NID:q190292; PIDN:AAA60160.1; PID:9553619
 A:Accession: JC5450
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 36-86, 'P', 88-122 <CHI2>
 A:Cross-references: GB:M68906; NID:q190290; PIDN:AAA60159.1; PID:9553618
 R:Tan, A.; To, H.P.
 submitted to Genbank, October 1995
 A:Reference number: A59269
 A:Accession: A59269
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-129 <TRAN>
 A:Cross-references: GB:U37529; NID:q1017792; PIDN:AAA79195.1; PID:q1017793
 R:Lat, J.P.; Douglas, S.D.; Rappaport, E.; Wu, J.M.; Ho, W.Z.
 submitted to Genbank, February 1998
 A:Description: Identification of a delta isoform of preprotachykinin mRNA in human mono
 A:Reference number: A59270
 A:Accession: A59270
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 36-96, 'W', 116-118 <LA11>
 A:Cross-references: GB:AF050656; NID:q3098594; PIDN:AA15702.1; PID:q3098595
 A:Experimental source: alpha splice form; tissue blood; tissue brain; cell type monocyte
 A:Accession: B59270
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 36-73,89-96, 'W', 116-122 <LA12>
 A:Cross-references: GB:AF050658; NID:q3098598; PIDN:AA15704.1; PID:q3098599
 A:Experimental source: delta splice form; tissue blood; tissue brain; cell type monocyte
 C:Comment: This protein is processed to produce the tachykinin peptide hormones neurokin
 in K).
 C:Genetics:
 A:Gene: GDB:TAC1; TAC2; NKNA; PPT-A
 A:Cross-references: GDB:119452; OMIM:162320
 A:Map position: 7q21-q22
 C:Superfamily: substance P precursor
 C:Keywords: alternative splicing; amidated carboxyl end; hormone; neuropeptide; tachykin
 F:1-129/Product: neurokinin 1 precursor, beta splice form #status predicted <SPB>
 F:1-96, 'W', 116-118/Product: neurokinin 1 precursor, alpha splice form #status predicted
 F:1-73,89-129/Product: neurokinin 1 precursor, gamma splice form #status predicted <SPG>
 F:1-73,89-96, 'W', 116-122/Product: neurokinin 1 precursor, alpha splice form #status pred
 F:1-19/Domin: signal sequence #status predicted <SIG>
 F:20-57/Domin: amino-terminal propeptide #status predicted <PRO>
 F:58-68/Product: neurokinin 1 #status experimental <NK1>
 F:72-107/Product: neuropeptide K #status predicted <NEK>
 F:98-107/Product: neurokinin 2 #status predicted <NK2>
 F:100-107/Product: neurokinin 2(3-10) #status experimental <NK23>
 F:101-107/Product: neurokinin 2(4-10) #status experimental <NK24>
 F:111-126/Domin: carboxyl-terminal propeptide #status experimental <CTP>
 F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following g1
 F:107/Modified site: amidated carboxyl end (Met) (amide in mature form from following g1

Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 86 FVGLMGKR 93
 |||||||
 Db 103 FVGLMGKR 110
 RESULT 12
 SPRTB
 substance P beta precursor - rat
 M:Alternate names: preprotachykinin beta; preprotachykinin gamma; substance K
 M:Contains: neurokinin A; substance P; substance P gamma precursor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1988 #sequence_revision 26-May-1995 #text_change 18-Jun-1999
 C:Accession: A37163; A26590; C26590; A25067; JC2411
 R:Carter, M.S.; Krause, J.E.
 J. Neurosci. 10, 2203-2214, 1990
 A:Title: Structure, expression, and some regulatory mechanisms of the rat preprotachy
 A:Reference number: A37163; MUID:90331040
 A:Accession: A37163
 A:Molecule type: DNA
 A:Residues: 1-130 <CAN>
 A:Cross-references: GB:M34159; GB:M34160; GB:M34162; NID:q206334; PIDN:AAA41926.1; PI
 R:Krause, J.E.; Chirgwin, J.M.; Carter, M.S.; Xu, Z.S.; Hershey, A.D.
 Proc. Natl. Acad. Sci. U.S.A. 84, 881-885, 1987
 A:Title: Three rat preprotachykinin mRNAs encode the neuropeptides substance P and ne
 A:Reference number: A34187; MUID:87118268
 A:Accession: A26590
 A:Molecule type: mRNA
 A:Residues: 1-130 <KAN>
 A:Cross-references: GB:M15191; NID:q206341; PIDN:AAA41928.1; PID:q206342; GB:M35277
 A:Accession: C26590
 A:Molecule type: mRNA
 A:Residues: 1-73,89-130 <KR2>
 A:Cross-references: GB:M34183; NID:q206343; PIDN:AAA41929.1; PID:q206344
 R:Kawaguchi, Y.; Hoshimaru, M.; Nawa, H.; Nakatsuki, S.
 Biochem. Biophys. Res. Commun. 139, 1040-1046, 1986
 A:Title: Sequence analysis of cloned cDNA for rat substance P precursor: existence of
 A:Reference number: A25067; MUID:87025808
 A:Accession: A25067
 A:Molecule type: mRNA
 A:Residues: 1-73,89-130 <KAN>
 A:Cross-references: GB:M14312; NID:q206339; PIDN:AAA41927.1; PID:q206340
 R:Khan, I.; Collins, S.M.
 Biochem. Biophys. Res. Commun. 202, 796-802, 1994
 A:Title: Fourth isoform of preprotachykinin messenger RNA encoding for substance P in
 A:Reference number: JC2411; MUID:94324969
 A:Accession: JC2411
 A:Molecule type: mRNA
 A:Residues: 48-110 <KHA>
 A:Experimental source: Intestine
 C:Comment: Alternative splicing of the mRNA for substance P precursor yields the beta
 C:Comment: The beta and gamma forms are processed to yield substance P and neurokinin
 C:Genetics:
 A:introns: 41/3; 74/1; 89/1; 97/1; 115/1
 C:Superfamily: substance P precursor
 C:Keywords: alternative splicing; amidated carboxyl end; hormone; neuropeptide; tachy
 F:1-130/Product: substance P beta precursor #status predicted <PREB>
 F:1-73,89-130/Product: substance P gamma precursor #status predicted <PREG>
 F:1-15/Domin: signal sequence #status predicted <SIG>
 F:58-68/Product: neurokinin A #status predicted <NKA>
 F:98-107/Product: neurokinin A #status predicted <NKA>
 F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following
 F:107/Modified site: amidated carboxyl end (Met) (amide in mature form from following

Query Match 6.6%; Score 8; DB 1; Length 130;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 86 FVGLMGKR 93
 |||||||

Db 103 FVGLMGKR 110

RESULT 13

SPBOH

neurokinin 1 precursor, beta splice form [validated] - bovine

N:Alternate names: neurokinin A; preprotachykinin; substance K; substance P

N:Contains: neurokinin 1; neurokinin 1 precursor, alpha splice form; neurokinin 1 precu

C:Species: Bos primigenius taurus (cattle)

C:Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 16-Jun-2000

C:Accession: A05093; A01557; A01557; B25067; A61460; J05454; I45966

R:Nawa, H.; Kotani, H.; Nakanishi, S.

Nature 312, 729-734, 1984

A:Title: Tissue-specific generation of two preprotachykinin mRNAs from one gene by alter

A:Reference number: A05093; MUID:85086245

A:Accession: A05093

A:Molecule type: DNA

A:Residues: 1-130 <NMA1>

A:Cross-references: GB:X02351; GB:M14786; NID:9655; PIDN:CAA26206.1; PID:g1197197

R:Nawa, H.; Hirose, T.; Takashima, H.; Inayama, S.; Nakanishi, S.

Nature 306, 32-36, 1983

A:Title: Nucleotide sequences of cloned cDNAs for two types of bovine brain substance P

A:Reference number: A93318; MUID:84039802

A:Accession: A01559

A:Molecule type: mRNA

A:Residues: 1-130 <NMA2>

A:Cross-references: GB:X00075; NID:g758; PIDN:CAA24939.1; PID:g759

A:Accession: A01557

A:Molecule type: mRNA

A:Residues: 1-96, 'M', 116-130 <NMA3>

A:Cross-references: GB:X00076; NID:g762; PIDN:CAA24942.1; PID:g763

Biochem. Biophys. Res. Commun. 139, 1040-1046, 1986

A:Title: Sequence analysis of cloned cDNA for rat substance P precursor: existence of a

A:Reference number: A25067; MUID:87025808

A:Accession: B25067

A:Molecule type: mRNA

A:Residues: 1-73, 89-130 <KAM>

R:McGregor, G.P.; Kage, R.; Thim, L.; Conlon, J.M.

J. Neurochem. 53, 1871-1877, 1989

A:Title: Quantitation and characterization of peptides from the C-terminal flanking regi

A:Reference number: A61460; MUID:90039314

A:Accession: A61460

A:Molecule type: protein

A:Residues: 111-126 <MCG>

A:Experimental source: Corpus striatum

R:Chiwakata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schulze, W.; Ivell, R.

Endocrinology 128, 2441-2448, 1991

A:Title: Tachykinin (substance-P) gene expression in Leydig cells of the human and mouse

A:Reference number: J05450; MUID:91209287

A:Accession: J05454

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 36-120, 'A', 122 <CHI>

A:Cross-references: GB:M68911; NID:g163591; PIDN:AAA30724.1; PID:g552335

C:Comment: The protein is processed to produce neurokinin 1 (substance P) and neurokinin

C:Genetics:

A:Gene: PPT-A

A:Introns: 41/3; 74/1; 89/1; 97/1; 115/1

C:Superfamily: substance P precursor

C:Keywords: alternative splicing; amidated carboxyl end; hormone; neuropeptide; tachykin

F:1-130/Product: neurokinin 1 precursor, beta splice form #status predicted <SP>

F:1-73, 89-130/Product: neurokinin 1 precursor, alpha splice form #status predicted <SP>

Query Match

Best Local Similarity 6.6%; Score 8; DB 1; Length 130;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93

Db 103 FVGLMGKR 110

RESULT 14

S47038

tachykinin 1 precursor - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999

C:Accession: S47038

R:Heitland, A.; Krueffer, M.; Juergen Maegert, H.J.; Forssmann, W.G.

submitted to the EMBL Data Library, July 1994

A:Reference number: S47038

A:Accession: S47038

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-130 <HEI>

A:Cross-references: EMBL:X80662; NID:g520917; PIDN:CAA56691.1; PID:g520918

C:Superfamily: substance P precursor

Query Match

Best Local Similarity 6.6%; Score 8; DB 2; Length 130;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93

Db 103 FVGLMGKR 110

RESULT 15

I52526

neurokinin 1 precursor - mouse

N:Alternate names: neurokinin A; preprotachykinin; substance K; substance P

N:Contains: neurokinin 1; neurokinin 2

C:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000

C:Accession: I52526; J05452; I62741

R:Kako, K.; Muneakata, E.; Hosaka, M.; Murakami, K.; Nakayama, K.

Biomed. Res. 14, 253-259, 1993

A:Title: Cloning and sequence analysis of mouse cDNAs encoding preprotachykinin A and

A:Reference number: I52526

A:Accession: I52526

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-130 <KAB>

A:Cross-references: GB:DI7584; NID:g407345; PIDN:BA04508.1; PID:g435121

R:Chiwakata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schulze, W.; Ivell, R.

Endocrinology 128, 2441-2448, 1991

A:Title: Tachykinin (substance-P) gene expression in Leydig cells of the human and mo

A:Reference number: J05450; MUID:91209287

A:Accession: J05452

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 36-122 <CHI>

A:Cross-references: GB:M68908; NID:g200467; PIDN:AAA39969.1; PID:g554260

C:Genetics:

A:Gene: PPT-A

C:Superfamily: substance P precursor

C:Keywords: amidated carboxyl end

F:1-19/Product: signal sequence #status predicted <SIG>

F:1-73, 89-130/Product: neurokinin 1 precursor, alpha splice form #status predicted <SP>

F:1-19/Product: neurokinin 1 precursor, gamma splice form #status predicted <SP>

F:20-57/Product: amino-terminal propeptide #status predicted <PRO>

F:58-68/Product: neurokinin 1 #status predicted <NK1>

F:98-107/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: neurokinin 2 #status predicted <NK2>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

F:111-126/Product: carboxyl-terminal propeptide #status predicted <CTP>

Query Match 6.6%; Score 8; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 FVGLMGKR 93
|||||||
DB 103 FVGLMGKR 110

Search completed: May 24, 2002, 17:06:52
Job time: 116 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 17:04:26 ; Search time 13.01 Seconds
(without alignments)
227.171 Million cell updates/sec

Title: US-09-852-659-68

Perfect score: 121

Sequence: 1 MRIMLFTALAFSLAQSG.....DVNOENVPSGILKYPRAE 121

Best Local Similarity 100.0%; Pred. No. 3.3e-92;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRIMLFTAILAFSLAOSGAVCKEPOEEVVPGGGSKRDPDLYQLORLFKSHSSLEGL 60
DB 1 MRIMLFTAILAFSLAOSGAVCKEPOEEVVPGGGSKRDPDLYQLORLFKSHSSLEGL 60

OY 61 LKALSOASTDPKESTSPKRDMDHDFVGLMGKRSVOPDSPT 101
DB 61 LKALSOASTDPKESTSPKRDMDHDFVGLMGKRSVOPDSPT 101

RESULT 2

US-09-215-096-1
; Sequence 1, Application US/09215096
; Patent No. 6008194

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Lal, Preeti

TITLE OF INVENTION: HUMAN PREPROTACHYKININ B

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/215,096

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/879,995

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0326 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 122 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: BRAITUT03

CLONE: 2109906

US-09-215-096-1

Query Match

Best Local Similarity 100.0%; Pred. No. 3.3e-92;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRIMLFTAILAFSLAOSGAVCKEPOEEVVPGGGSKRDPDLYQLORLFKSHSSLEGL 60

DB 1 MRIMLFTAILAFSLAOSGAVCKEPOEEVVPGGGSKRDPDLYQLORLFKSHSSLEGL 60

OY 61 LKALSOASTDPKESTSPKRDMDHDFVGLMGKRSVOPDSPT 101

DB 61 LKALSOASTDPKESTSPKRDMDHDFVGLMGKRSVOPDSPT 101

RESULT 3

US-08-879-995A-4

; Sequence 4, Application US/08879995A

; Patent No. 5985606

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Kaser, Matthew R.

TITLE OF INVENTION: HUMAN PREPROTACHYKININ B

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/879,995A

FILING DATE: Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0326 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 116 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 205725

US-08-879-995A-4

Query Match

Best Local Similarity 12.4%; Score 15; DB 2; Length 116;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 KRDMHDFVGLMGKR 93

DB 80 KRDMHDFVGLMGKR 94

US-09-215-096-4

; Sequence 4, Application US/09215096

; Patent No. 6008194

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Lal, Preeti

TITLE OF INVENTION: HUMAN PREPROTACHYKININ B

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/215,096
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/879,995
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0326 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 205725
US-09-215-096-4

Query Match 12.4%; Score 15; DB 3; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 KRDMHDFVGLMKR 93
|||
DB 80 KRDMHDFVGLMKR 94

RESULT 5
US-08-879-995A-3
Sequence 3, Application US/08879995A
Patent No. 5985606
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,995A
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0326 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 163590
US-08-879-995A-3

Query Match 12.4%; Score 15; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 KRDMHDFVGLMKR 93
|||
DB 84 KRDMHDFVGLMKR 98

RESULT 6
US-09-215-096-3
Sequence 3, Application US/09215096,
Patent No. 6008194
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/215,096
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/879,995
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0326 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:

LIBRARY: GenBank
CLONE: 163590
us-09-215-096-3

Query Match 12.4%; Score 15; DB 3; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 KRDHDFVGLMKR 93
|||||
DB 84 KRDHDFVGLMKR 98

RESULT 7
US-07-899-205-3
; Sequence 3, Application US/07899205
; Patent No. 5288730
; GENERAL INFORMATION:
; APPLICANT: Baker, Raymond
; APPLICANT: Teall, Martin R.
; APPLICANT: Swain, Christopher J.
; APPLICANT: Williams, Brian J.
; TITLE OF INVENTION: AZABICYCLIC COMPOUNDS PHARMACEUTICAL
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 E. Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/899,205
; FILING DATE: 19920616
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Polk, Manfred
; REGISTRATION NUMBER: 27,102
; REFERENCE/DOCKET NUMBER: T-1106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-4285
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-899-205-3

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 81 DMHDFVGLM 90
|||||
DB 1 DMHDFVGLM 10

RESULT 8
US-08-184-935-10
; Sequence 10, Application US/08184935
; Patent No. 5476770
; GENERAL INFORMATION:

APPLICANT: PRADELLES, PHILIPPE
; TITLE OF INVENTION: IMMUNOMETRIC DETERMINATION OF AN ANTIGEN
; OR HAPTEN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBION, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,935
; FILING DATE: 24-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBION, NO. 5476770man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 846-286-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note="C-terminal amide"
; US-08-184-935-10

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 81 DMHDFVGLM 90
|||||
DB 1 DMHDFVGLM 10

RESULT 9
US-08-269-288-4
; Sequence 4, Application US/08269288
; Patent No. 5461140
; GENERAL INFORMATION:
; APPLICANT: Bruns, Robert F.
; APPLICANT: Gehlert, Donald R.
; APPLICANT: Howbert, James J.
; APPLICANT: Lupp, William H.W.
; TITLE OF INVENTION: NAPHTHYL TACHYKININ RECEPTOR ANTAGONISTS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center/1104
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/269,288
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-9715
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 276-0756
TELEFAX: (317) 276-3861
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-269-288-4

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred.No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 81 DMHDFVGLM 90
DB 1 DMHDFVGLM 10

RESULT 10
US-08-338-484-3
Sequence 3, Application US/08338484
Patent No. 5494926
GENERAL INFORMATION:
APPLICANT: Owens, Andrew P.
APPLICANT: Teall, Martin R.
APPLICANT: Williams, Brian J.
TITLE OF INVENTION: 2/3-(HETEROCYCLIC ALKYL
TITLE OF INVENTION: AMINO)-1-(SUBSTITUTED PHENYL-METHOXY)-ETHANES/PROPANES AS
TITLE OF INVENTION: TACHYKININ RECEPTOR ANTAGONISTS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. Robert J. No. 5494926th
STREET: 126 E. Lincoln Ave., P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,484
FILING DATE: 18-NOV-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: No. 5494926th, Robert J.
REGISTRATION NUMBER: 27,366
REFERENCE/DOCKET NUMBER: T-1158
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4720
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-338-484-3

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred.No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 81 DMHDFVGLM 90
DB 1 DMHDFVGLM 10

RESULT 11
US-08-175-432-3
Sequence 3, Application US/08175432
Patent No. 5495047
GENERAL INFORMATION:
APPLICANT: Saari, Walfrid S.
APPLICANT: van Niel, Monique B.
APPLICANT: Williams, Brian J.
TITLE OF INVENTION: FUSED TRICYCLIC COMPOUNDS,
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR USE
TITLE OF INVENTION: IN THERAPY
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: NORTH, ROBERT J.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,432
FILING DATE: 07-JAN-1994
CLASSIFICATION: 560
ATTORNEY/AGENT INFORMATION:
NAME: No. 5495047th, Robert J.
REGISTRATION NUMBER: 27,366
REFERENCE/DOCKET NUMBER: T-1152Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-7262
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-175-432-3

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred.No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 81 DMHDFVGLM 90
DB 1 DMHDFVGLM 10

RESULT 12
US-08-462-413-4
Sequence 4, Application US/08462413
Patent No. 553009
GENERAL INFORMATION:

APPLICANT: Cho, Sung Y.
APPLICANT: Copp, James D.
APPLICANT: Glnah, Francis O.
APPLICANT: Hansen, Guy J.
APPLICANT: Hipskind, Phillip A.
APPLICANT: Huff, Bret E.
APPLICANT: Martinehl, Michael J.
APPLICANT: Staszak, Michael A.
APPLICANT: Tharp-Taylor, Roger W.
TITLE OF INVENTION: PROCESS FOR PREPARING NON-PEPTIDYL
TITLE OF INVENTION: TACHIKININ RECEPTOR ANTAGONISTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States of America
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,413
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/271,708
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-9475
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 276-0756
TELEFAX: (317) 276-3861
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-462-413-4

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 DMHDFVGLM 90
DB 1 DMHDFVGLM 10

RESULT 13
US-08-391-910-4
Sequence 4, Application US/08391910
Patent No. 556313
GENERAL INFORMATION:
APPLICANT: Hipskind, Phillip A.
TITLE OF INVENTION: HEXAMETHYLENIMINYL TACHIKININ RECEPTOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States of America
ZIP: 46285

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,910
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-9979
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 276-0756
TELEFAX: (317) 276-3861
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-391-910-4

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 DMHDFVGLM 90
DB 1 DMHDFVGLM 10

RESULT 14
US-08-418-994-4
Sequence 4, Application US/08418994
Patent No. 5565568
GENERAL INFORMATION:
APPLICANT: Cho, Sung-Yong S.
APPLICANT: Hipskind, Phillip A.
APPLICANT: Howbert, J. J.
APPLICANT: Muehl, Brian S.
APPLICANT: Nixon, James A.
TITLE OF INVENTION: 2-ACETAMINOPROPANAMIDES AS TACHIKININ
TITLE OF INVENTION: RECEPTOR ANTAGONISTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States of America
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/418,994
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-8252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 276-0756
TELEFAX: (317) 276-3861
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-418-994-4

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 DMHDFVGLM 90
|||||
Db 1 DMHDFVGLM 10

RESULT 15

US-08-391-814-4
Sequence 4, Application US/08391814
Patent No. 5607947
GENERAL INFORMATION:
APPLICANT: Hipskind, Philip A.
TITLE OF INVENTION: PYROLIDINYL TACHYKININ RECEPTOR
TITLE OF INVENTION: ANTAGONISTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States of America
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,814
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-9965
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 276-0756
TELEFAX: (317) 276-3861
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-391-814-4

Query Match 8.3%; Score 10; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 DMHDFVGLM 90
|||||
Db 1 DMHDFVGLM 10

Search completed: May 24, 2002, 17:06:29
Job time: 123 sec

PR 06-JUN-1997: 97US-0048970.
 PR 05-SEP-1997: 97US-0057765.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ferrle AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;
 PI Li H, Moore PA, Rosen CA, Ruben SM, Soppet DR;
 PI Wei YF, Young PE, Zeng Z;
 XX
 DR WPI: 1998-520811/44.
 DR N-PSDB: AAV34302.
 XX
 PT Isolated human poly:nucleotide(s) encoding secretory peptide(s) -
 PT used to develop products for the diagnosis and treatment of e.g.
 PT inflammation, cancers, CNS disorders or immune system disorders
 XX
 PS Claim 1: Page 162-163; 201pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AAV34277) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 28 novel genes and their fragments (nucleic
 CC acid sequences: AAV34286-V34325; amino acid sequences AAV75196-W75235)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 28
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV34286 for described uses).
 CC
 SQ Sequence 121 AA:

Query Match 100.0%; Score 121; DB 19; Length 121;
 Best Local Similarity 100.0%; Pred. No. 1.3e-111;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRIMLFTIILAFSLAOSGAVCKEPOEYVGGGSKRDPDLYOLOLRFKSHSLEGL 60
 DB 1 mrimlftailafslagsfgavckepgeevp99gskrdpdllyqlilfkshslegl 60
 QY 61 LKALSOASTDPKESISPEKRDHDFVGLMGKRSYVQDSPTDVNOENVPFGLKYPRA 120
 DB 61 lkalsqastdpkesistpekrdmhdfvglmgrksyvpdsptdvnoenvpfglkyprra 120
 QY 121 E 121
 DB 121 e 121

RESULT 2
 AAM97213
 ID AAM97213 standard; Protein: 121 AA.
 XX
 AC AAM97213;
 XX
 DT 07-MAY-1999 (first entry)
 XX
 DE A human zneurok1 polypeptide.
 XX
 XX Human; zneurok1; neurokinin B; prohormone convertase; cell growth;
 KW modulation; inflammation; nonreception; emesis; contraction;
 KW hormone secretion; DNA synthesis; inositol phosphate turnover;
 KW arachidonate release; phospholipase-C activation; gastric emptying;
 KW human neutrophil activation; ADCC capability;
 KW superoxide anion production; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN MO9855612-A1.

XX
 PD 10-DEC-1998.
 XX
 PF 28-MAY-1998: 98WO-US10842.
 XX
 PR 02-JUN-1997: 97US-0048290.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Sheppard PO;
 PI
 DR WPI: 1999-070268/06.
 DR N-PSDB: AAX15447.
 XX
 PT New isolated neurokinin polypeptides, zneurok1 - used to develop
 PT products for modulating e.g. inflammation, nonreception, emesis,
 PT muscle contraction, hormone secretion, DNA synthesis or cell growth
 XX
 PS Claim 3: Page 76-77; 100pp; English.
 XX
 CC The present sequence represents a human zneurok1 polypeptide. The
 CC polypeptide releases a neurokinin B polypeptide in the presence
 CC of a prohormone convertase capable of cleaving dibasic amino
 CC acids. The zneurok1 polypeptides can be used for modulating inflammation,
 CC nonreception or emesis. The polypeptides, fragments, fusion proteins,
 CC agonists, antagonists or antibodies may also modulate contraction,
 CC hormone secretion, DNA synthesis or cell growth, inositol phosphate
 CC turnover, arachidonate release, phospholipase-C activation, gastric
 CC emptying, human neutrophil activation or ADCC capability, or superoxide
 CC anion production. The polynucleotides can also be used for gene therapy.
 CC The products can also be used for detection, diagnosis and screening
 CC assays.
 CC
 SQ Sequence 121 AA:

Query Match 100.0%; Score 121; DB 20; Length 121;
 Best Local Similarity 100.0%; Pred. No. 1.3e-111;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRIMLFTIILAFSLAOSGAVCKEPOEYVGGGSKRDPDLYOLOLRFKSHSLEGL 60
 DB 1 mrimlftailafslagsfgavckepgeevp99gskrdpdllyqlilfkshslegl 60
 QY 61 LKALSOASTDPKESISPEKRDHDFVGLMGKRSYVQDSPTDVNOENVPFGLKYPRA 120
 DB 61 lkalsqastdpkesistpekrdmhdfvglmgrksyvpdsptdvnoenvpfglkyprra 120
 QY 121 E 121
 DB 121 e 121

RESULT 3
 AAM74413
 ID AAM74413 standard; Protein: 121 AA.
 XX
 AC AAM74413;
 XX
 DT 10-MAY-1999 (first entry)
 XX
 DE HPMBQ91 protein sequence.
 XX
 XX HPMBQ91; neurokinin B precursor; neuronal disorder; CNS related disorder;
 KW gastrointestinal disorder; cardiovascular disorder; metabolic disorder;
 KW smooth muscle disorder; inflammatory disorder; cancer; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP892053-A2.
 XX
 PD 20-JAN-1999.


```

PE 26-JUN-1998; 98EP-0305066.
XX
PR 14-JUL-1997; 97EP-0305215.
XX
PA (HOMA-) HUMAN GENOME SCI.
XX (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI Duckworth DM, Hastings GA, Ruben SM;
XX
PI WPI: 1999-083570/08.
XX N-PSDB; AAX18197.
XX
PT New human neurokinin B precursor (HPMB091) polypeptides and
PT polynucleotide - useful as diagnostic reagents and for prevention
PT and treatment of neuronal, metabolic, inflammatory and
PT gastrointestinal disorders, and cancers
XX
PS Claim 11; Page 14-15; 18pp; English.
XX
XX This sequence is the human neurokinin B precursor HPMB091 of the
CC invention. HPMB091 polypeptides and polynucleotides are useful for
CC diagnosing susceptibility to diseases associated with HPMB091 protein
CC imbalance by determining HPMB091 polypeptide expression levels. HPMB091
CC polypeptides can be used to screen for agonists and antagonists by
CC measuring the binding to HPMB091, and observing the stimulation or
CC inhibition of HPMB091 function. These can be used in treatment to
CC activate or inhibit HPMB091 activity to treat conditions associated with
CC a lack of HPMB091 protein. Gene therapy may also be used to affect
CC endogenous HPMB091 polypeptide expression. HPMB091 antibodies are useful
CC for inducing an immune response to immunise and prevent disease, and for
CC isolating HPMB091 clones or purifying the polypeptides by affinity
CC chromatography. HPMB091 polypeptides can be administered directly or as a
CC vaccine to inoculate against disease. Diseases diagnosed, prevented and
CC treated include: neuronal disorders; CNS related disorders;
CC gastrointestinal and cardiovascular disorders; metabolic disorders
CC including diabetes and obesity; smooth muscle disorders; inflammatory
CC disorders; and cancers including adenomas, leiomyomas, liposarcomas,
CC melanomas, pulmonary chondroid hamartomas, lung, prostate and breast
CC cancer. The HPMB091 polypeptide is also useful for mapping the gene to a
CC chromosome, allowing gene inheritance to be studied through linkage
CC analysis, and tissue localisation studies, for determining HPMB091
CC expression patterns.
XX
SO Sequence 121 AA:

Query Match 100.0%; Score 121; DB 20; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.3e-111;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MIMIMLFITILFSLAQSGAVCKEFGEEVYFEGGSKRDPDLYLLQRLFKSHSLEGL 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 mIMiIlfItIlLlFsLagSfgavckEpgEevYvPggsgrkrdpdylllqrLfkshsSlegl 60

QY 61 LKALQASQSDPPESTSPKRDHMDPFVGLMGKRSVOPDSDPTVONQNVSPFGLTKYPPRA 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 lKaLqasqsdPkesstspEkrdmhdfvGlmgkRsvOpdSdptvOnqnvSpfGlLkYppra 120

QY 121 E 121
   |
DB 121 e 121

RESULT 4
ID AAB82380
XX AAB82380 standard; protein; 121 AA.
XX
XX AAB82380;
XX
XX 23-JUL-2001 (first entry)
XX
XX Human neurokinin B precursor.
XX

```

KW Neurokinin B: human; pregnancy; hypertension; pre-eclampsia;
KW diagnosis; therapy.

XX Homo sapiens.
OS
XX Key Location/Qualifiers
XX FH 81..91
XX Peptide /label= Neurokinin
FT
FN WO200136979-A2.
PN
XX
XX 25-MAY-2001.
PD
XX
XX 10-NOV-2000; 2000WO-GB04315.
PF
XX 16-NOV-1999; 99GB-0027125.
PR
XX (UTRE-) UNIV READING.
PA
XX Page N, Lowry P;
PI
XX WPI: 2001-355676/37.
DR N-PSDB; AAF90333; AAF90334.
XX

PT Detecting production of the human precursor of neurokinin B by the
FT placenta in pregnancy induced hypertension or pre-eclampsia or related
PT foetal complication
XX
XX Example 1; Fig 1; 63pp; English.
PS

CC The present sequence is that of human neurokinin B (NKB) precursor.
CC The cloning of placental cDNA (see AAF90333) was used to identify
CC the NKB precursor. The precursor is processed to the 10-amino acid
CC NKB peptide. Detection of raised plasma levels of NKB, NKB
CC precursor, its breakdown product or variants at an early stage of
CC pregnancy provide an indication of the likely development of
CC pregnancy induced hypertension or pre-eclampsia. Reduction in the
CC levels of circulating NKB (or reduction of its effects) will
CC ameliorate the adverse effects upon the mother seen in these
CC conditions. Thus, the invention provides methods for predicting or
CC diagnosing pregnancy induced hypertension, pre-eclampsia or
CC related foetal complication based on measuring NKB levels in the
CC blood, and methods for preventing or treating these conditions,
CC e.g. by administering an agent that inhibits the biological
CC effect of NKB, such as an NK1, NK2 or NK3 antagonist.
CX
SX Sequence 121 AA;

SQ

Query Match 100.0%; Score 121; DB 22; Length 121;
Best Local Similarity 100.0%; Pred. NO. 1.3e-111; Indels 0; Gaps 0;
Matches 121; Conservative 0; Mismatches 0;

OY 1 NMIMLFTAILAFLASGFGAVCKEPOEVYVGGRSKRDPDLXQLRLFKSHSLBGL 60
Db 1 nmimlftailafslasgfgavckepgeevyvggrskrdpdlxqlrlfkshslegl 60

OY 61 LKALQSASTDPEKHSISPEKRDMDHDFVGLMGKRKSQVPDSPTDVNENWPSEILKTPPRA 120
Db 61 lkalsqastdpkhsispekrdmdhdfvlgmgkrksvqpsdptdvnenwpsfigilkyppra 120

OY 121 E 121
|
Db 121 e 121

RESULT 5
ID ABB11878
XX ABB11878 standard; peptide: 137 AA.
AC
XX ABB11878;
XX
DT 11-JAN-2002 (first entry)

XX DE Human neurokinin B-like protein homologue, SFG ID NO:2248.
 XX DE
 XX DE Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antistatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiac; virocidic; antibacterial;
 KW antifungal; vulnery; antitumor.
 XX OS Homo sapiens.
 XX PN MO200157188-A2.
 XX PD
 XX PD 09-AUG-2001.
 XX PF 05-FEB-2001, 2001WO-US03800.
 XX PR 03-FEB-2000; 2000US-0496914.
 XX PR 27-APR-2000; 2000US-0560875.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Dzmanac RT;
 XX DR MPI; 2001-457740/49.
 XX DR N-PSDB; ABA09122.
 XX PT Human proteins and DNA encoding sequences useful for preventing,
 XX PT treating or ameliorating a medical condition in a mammalian subject
 XX PT e.g. arthritis and cancer -
 XX PS
 XX PS Claim 20; Page 270; 1963pp; English.
 XX CC Sequences ABA010981-ABA01230 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells

CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 XX
 XX SQ Sequence 137 AA;
 XX
 XX Query Match 100.0%; Score 121; DB 22; Length 137;
 XX Best Local Similarity 100.0%; Pred. No. 1.4e-111;
 XX Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 MRLMLFTALIASLAOSFGAVCKEPEEYVPGGGRKRPDLYQLQRLFKSHSLEGL 60
 XX Db 17 mrlmlftalialslagsfivakepeeyvpgggrskrpdyqlqlfkshslegl 76
 XX
 XX QY 61 LKALSOASTDPKKESTPEKRDMDFFVGLMGKRSVOPDPTDVNOENVPSFGLKYPRA 120
 XX Db 77 lkalsqastdpkkestspekrmdmdffvglmgrksvopdpdptdvngenvpsfgllkyprra 136
 XX
 XX QY 121 E 121
 XX Db 137 e 137
 XX
 XX RESULT 6
 XX AAW96144
 XX ID AAW96144 standard; Protein; 122 AA.
 XX AC
 XX AAW96144;
 XX DT 27-APR-1999 (first entry)
 XX DE Human preprotachykinin B.
 XX KW
 XX KW Preprotachykinin B; PPT-B; neoplastic disorder;
 KW neurological disorder; Alzheimer's disease; amnesia;
 KW cerebral neoplasms; dementia; depression; Down's syndrome;
 KW Huntington's disease; multiple sclerosis; Parkinson's disease;
 KW paranoid psychoses; schizophrenia; Tourette's disorder; angina;
 KW anaplastic shock; asthma; cardiovascular shock;
 KW myocardial infarction; migraine.
 XX OS Homo sapiens.
 XX FH
 XX Key Location/Qualifiers
 XX FT Misc-difference 104
 XX FT /Label= Leu, Ser or Trp
 XX PN
 XX PN WO9857986-A2.
 XX PD 23-DEC-1998.
 XX PF 19-JUN-1998; 98WO-US12855.
 XX PR 19-JUN-1997; 97US-0879995.
 XX RA (INCY-) INCYTE PHARM INC.
 XX PI Hillman JJ, Kaser MR, Lal P;
 XX DR MPI; 1999-080948/07.
 XX DR N-PSDB; AAX08906.
 XX PT New human preprotachykinin B - useful for treating neurological
 XX PT disorders and cancer
 XX PS Claim 1; Page 48-49; 57pp; English.
 XX CC Human preprotachykinin B (PPT-B) can be used to treat a
 CC neurological disorder. Antagonists of PPT-B can also be used in
 CC the treatment of neoplastic disorders. Particular neurological,

CC disorders include akathisia, Alzheimer's disease, amnesia,
 CC amyotrophic lateral sclerosis, bipolar disorder, cataplexy,
 CC cerebral neoplasms, dementia, depression, Down's syndrome, tardive
 CC dyskinesia, dystonia, Huntington's disease, multiple sclerosis,
 CC neurofibromatosis, Parkinson's disease, paranoid psychosis,
 CC schizophrenia, and Tourette's disorder. PPT-B or its agonist can
 CC also be used to treat angina, anaphylactic shock, arrhythmias,
 CC asthma, cardiovascular shock, Cushing's syndrome, hypertension,
 CC hypoglycemia, myocardial infarction, migraine and pheochromocytoma.
 XX

Sequence 122 AA:

Query Match 83.5%; Score 101; DB 20; Length 122;
 Best Local Similarity 100.0%; Pred. No. 7, 3e-92;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRIMLFTALIAFSLAOSFGAVCKEPEQEVVPGGGRKRDPLYLQLRLFKSHSSLEGL 60
 Db 1 mrimlftalialslasfgavckepqevvpgggrskrdpllyqlrlfkshsslegl 60
 QY 61 LKALSOASTDPKSTSPKRDMDHDFVGLMGKRSVQPDPT 101
 Db 61 lkalsastdpkstspkrdmdhdfvglmgrsvqpdpt 101

RESULT 7

AAB33445
 ID AAB33445 standard; Protein: 135 AA.

AC AAB33445;

DE 29-JAN-2001 (first entry)

XX Human PRO1155 protein UNQ585 SEQ ID NO:157.

XX Human; immune related disease; diagnosis; antinflammatory; cardiac;
 KW dermatological; antiarthritic; antirheumatic; immunosuppressive;
 KW hemostatic; antihypertensive; antidiabetic; neuroprotective;
 KW antineoplastic; hepatotropic; virucide; antiparasitic; antiallergic;
 KW antiaesthetic; systemic lupus erythematosus; rheumatoid arthritis;
 KW osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;
 KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
 KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
 KW autoimmune thrombocytopenia; immune-mediated renal disease;
 KW demyelinating disease; hepatobiliary disease; Whipple's disease;
 KW inflammatory bowel disease; gluten-sensitive enteropathy;
 KW autoimmune disease; immune-mediated skin disease; allergic disease;
 KW immunological disease; transplantation associated disease;
 KW graft rejection; graft-versus-host-disease.

XX Homo sapiens.

PN WO200053758-A2.

XX 14-SEP-2000.

PF 02-MAR-2000; 2000WO-US05841.

XX 08-MAR-1999; 99WO-US05028.
 PR 10-MAR-1999; 99US-0123618.
 PR 12-MAR-1999; 99US-0123957.
 PR 23-MAR-1999; 99US-0125775.
 PR 12-APR-1999; 99US-0128849.
 PR 20-APR-1999; 99WO-US08615.
 PR 28-APR-1999; 99US-0134445.
 PR 04-MAY-1999; 99US-0133371.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 28-JUL-1999; 99US-0146222.

PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-OCT-1999; 99US-0162506.
 PR 29-NOV-1999; 99WO-US28214.
 PR 30-NOV-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 16-DEC-1999; 99WO-US28565.
 PR 16-DEC-1999; 99WO-US30095.
 PR 30-DEC-1999; 99WO-US30099.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 05-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 XX
 XX (GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;
 PI Kabakoff RC, Lu Y, Pan Y, Pennica D, Shelton DL, Smith V;
 PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;
 DR N-PSDB; AAC58610.
 XX WPI: 2000-572271/53.

XX Sixty four PRO polypeptides, useful in the diagnosis and treatment of
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
 XX Claim 33; Fig 64; 309pp; English.

XX The present invention describes sixty four human PRO proteins which can
 CC be used in the treatment of immune related diseases. The human PRO
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
 CC treating and diagnosing immune related disorders. The disorders are
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune hemolytic
 CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,
 CC immune-mediated renal disease, demyelinating diseases of the central
 CC and peripheral nervous systems, hepatobiliary diseases, inflammatory
 CC bowel disease, gluten-sensitive enteropathy and Whipple's disease,
 CC autoimmune or immune-mediated skin diseases, allergic diseases,
 CC immunological diseases of the lung, and transplantation associated
 CC diseases including graft rejection and graft-versus-host-disease.
 CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAC58579 to AAC58647 and
 CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.

XX Sequence 135 AA:

Query Match 80.2%; Score 97; DB 21; Length 135;
 Best Local Similarity 100.0%; Pred. No. 7e-88; Indels 0; Gaps 0;
 Matches 97; Conservative 0; Mismatches 0;

QY 1 MRIMLFTALIAFSLAOSFGAVCKEPEQEVVPGGGRKRDPLYLQLRLFKSHSSLEGL 60
 Db 1 mrimlftalialslasfgavckepqevvpgggrskrdpllyqlrlfkshsslegl 60
 QY 61 LKALSOASTDPKSTSPKRDMDHDFVGLMGKRSVQPDPT 97

Db 61 lkalsgastdpkestapekradmbdfvglmkrsvqp 97
RESULT 8
ID AAY66739 standard; protein; 135 AA.
AC AAY66739;
XX
DT 05-APR-2000 (first entry)
DE Membrane-bound protein PRO1155.
XX
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping.
XX
OS Homo sapiens.
XX
PN WO9963088-A2.
XX
PD 09-DEC-1999.
XX
PE 02-JUN-1999; 99WO-US12252.
XX
PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 03-JUN-1998; 98US-0087759.
PR 04-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088625.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 12-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 16-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 18-JUN-1998; 98US-0089601.
PR 18-JUN-1998; 98US-0089607.
PR 18-JUN-1998; 98US-0089608.
PR 19-JUN-1998; 98US-0089647.
PR 19-JUN-1998; 98US-0089948.

PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 24-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 25-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090680.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 26-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090697.
PR 01-JUL-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 10-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 11-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096756.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 19-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.

PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098525.
PR 16-SEP-1998; 98US-0100634.
PR 12-JAN-1999; 99US-0115565.
XX (GETH) GENENTECH INC.
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PI Wood WI, Yuan J;
XX WPI: 2000-072883/06.
DR N-PSDB; AAs65085.
XX
PT Membrane-bound proteins and related nucleotide sequences.
XX
PS Claim 12; Fig 254; 822pp; English.
XX
CC The invention provides membrane-bound PRO polypeptides and
CC polynucleotides encoding them. The PRO sequences of the invention were
CC identified based on extracellular domain homology screening. The PRO
CC sequences have homology with proteins including LDL receptors, TIE
CC ligands and various enzymes. The membrane-bound proteins and receptor
CC molecules are useful as pharmaceutical and diagnostic agents. Receptor
CC immunoadhesins, for instance, can be used as therapeutic agents to block
CC employed for screening of potential peptide or small molecule inhibitors
CC of the relevant receptor/ligand interaction. The PRO encoding sequences
CC are useful as hybridization probes, in chromosome and gene mapping and in
CC the generation of antisense RNA and DNA. PRO nucleic acid sequences
CC will also be useful for the preparation of PRO polypeptides, especially
CC by recombinant techniques.
XX
SQ Sequence 135 AA;
Query Match 80.2%; Score 97; DB 21; Length 135;
Best Local Similarity 100.0%; Pred. No. 7e-88;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRIMLFTALIAFSLAQSFGAVCKEPEVEVPGGGRKRDPDLYOLLQRLFKSHSLEGL 60
Db 1 mrimlftalialafslagsfgavckepveevpgggrskrdpdllyollqrlfkshslegl 60
QY 61 LKALISQASTDPKSTSPKSRDMHDFYGLMGKRSVQP 97
Db 61 lkalsqastdpkstsprkdmhdfyglmgkrsvqp 97
RESULT 9
AAU29245
ID AAU29245 standard; Protein; 135 AA.
XX
AC AAU29245;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #222.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

XX Homo sapiens.
OS
XX WO20016848-A2.
PN
XX
XX 20-SEP-2001.
PD
XX
XX 28-FEB-2001; 2001WO-US06520.
PF
XX
XX 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186368P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195375P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX
PA (GETH) GENENTECH INC.
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI: 2001-602746/68.
DR N-PSDB; AAs46146.
XX
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
XX presence of tumours, such as prostate and breast tumours, in mammals and
XX to screen for modulators of the compounds -
XX
XX Claim 11; Fig 444; 774pp; English.
XX
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
XX The PRO polypeptides and their associated nucleic acids can be used to
XX detect the presence of a tumour in a mammal by comparing the level of
XX expression of a PRO polypeptide in a test sample of cells from the animal
XX and a control sample of normal cells, whereby a higher level of
XX expression in the test sample indicates the presence of a tumour in the
XX mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
XX and rabbits but are preferably human. The polypeptides can be used to
XX stimulate tumour necrosis factor (TNF) alpha release from human blood,
XX when contacted with it. A specific polypeptide can be used to stimulate

CC the proliferation or differentiation of chondrocyte cells. The PRO
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours. In mammalian
 CC subjects, the oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.
 XX
 SQ Sequence 135 AA;

Query Match 80.2%; Score 97; DB 22; Length 135;
 Best Local Similarity 100.0%; Pred. No. 7e-88;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRIMLFTAILAFSLAOSFGAVCKEPOEYVPGGSRKRPDLYQLLRFKSHSIEGL 60
 DB 1 mrimlftailafslagsfgavckepgeevpgggrskrdpdllyqlrlfkshslegl 60
 QY 61 LKALSQASTDPKRETSPEKRDHDFVGLMGKRSVQP 97
 DB 61 lkalsqastdpkrestspekrmdhdfvglmgrsvqp 97

RESULT 10
 AAB65262
 ID AAB65262 standard; Protein; 135 AA.
 AC AAB65262;
 XX

DT 02-APR-2001 (first entry)
 XX

DE Human PRO1155 (UNQ585) protein sequence SEQ ID NO:359.

XX Human; secreted and transmembrane protein; PRO; cytosolic;
 KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
 KM diagnostic assay.
 XX

OS Homo sapiens.

PN WO200073454-A1.

PD 07-DEC-2000.

PE 30-MAR-2000; 2000MO-US08439.

XX 02-JUN-1999; 99MO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 07-JUL-1999; 99US-0143048.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 28-JUL-1999; 99US-0146222.
 PR 17-AUG-1999; 99US-0149396.
 PR 15-SEP-1999; 99US-0158663.
 PR 15-SEP-1999; 99US-021090.
 PR 08-OCT-1999; 99US-021547.
 PR 30-NOV-1999; 99MO-US28313.
 PR 01-DEC-1999; 99MO-US28313.
 PR 16-DEC-1999; 99MO-US30095.
 PR 20-DEC-1999; 99MO-US30911.
 PR 05-JAN-2000; 2000MO-US00219.
 PR 06-JAN-2000; 2000MO-US00376.
 PR 11-FEB-2000; 2000MO-US03565.
 PR 18-FEB-2000; 2000MO-US04341.
 PR 22-FEB-2000; 2000MO-US04414.
 PR 24-FEB-2000; 2000MO-US04914.
 PR 02-MAR-2000; 2000MO-US05004.
 PR 15-MAR-2000; 2000MO-US05841.
 PR 20-MAR-2000; 2000MO-US06884.
 XX (GETH) GENENTECH INC.
 PA Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL.

PI Ferrara N, Fong S, Gerner H, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi CJ, Gurney AL, Kijavrin TJ, Napier MA, Pan J, Paoletti NF;
 PI Roy MA, Stewart JA, Tumas D, Watanabe CK, Williams PM, Wood WI;
 PI Zhang Z;

DR WPI: 2001-032160/04.
 DR N-PSDB; AAF44231.

PT PRO polynucleotides used to produce polypeptides used to target
 PT bioactive molecules such as toxins, radiolabels or antibodies, to
 PT specific cells, to cause targeted cell death -
 XX
 PS Claim 12; Fig 254; 935pp; English.

CC The present invention describes human secreted and transmembrane PRO
 CC proteins. The PRO proteins have cytosolic activity. The PRO proteins
 CC can be used for targeted delivery of bioactive molecules, such as
 CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
 CC sequences, and their fragments, can be used as hybridisation probes. In
 CC chromosomal and gene mapping, and in the generation of anti-gene RNA
 CC and DNA. They may also be used to produce transgenic animals which are
 CC used to develop and screen therapeutically useful reagents. The PRO
 CC nucleotide and protein sequence can be used for tissue typing and in
 CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
 CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
 CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.

SQ Sequence 135 AA;

Query Match 80.2%; Score 97; DB 22; Length 135;
 Best Local Similarity 100.0%; Pred. No. 7e-88;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRIMLFTAILAFSLAOSFGAVCKEPOEYVPGGSRKRPDLYQLLRFKSHSIEGL 60
 DB 1 mrimlftailafslagsfgavckepgeevpgggrskrdpdllyqlrlfkshslegl 60
 QY 61 LKALSQASTDPKRETSPEKRDHDFVGLMGKRSVQP 97
 DB 61 lkalsqastdpkrestspekrmdhdfvglmgrsvqp 97

RESULT 11
 AAW75228
 ID AAW75228 standard; Protein; 121 AA.

AC AAW75228;
 XX

DT 29-JAN-1999 (first entry)
 XX

DE Human secreted protein encoded by gene 17 clone HPM091.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 67 /Label- unknown

FT Misc-difference 89 /Label- unknown

PN WO9840483-A2.

```

XX 17-SEP-1998.
XX
XX 12-MAR-1998; 98WO-US04858.
XX
PR 19-DEC-1997; 97US-0068368.
PR 14-MAR-1997; 97US-0040710.
PR 14-MAR-1997; 97US-0040762.
PR 30-MAY-1997; 97US-0048100.
PR 30-MAY-1997; 97US-0048189.
PR 30-MAY-1997; 97US-0048357.
PR 30-MAY-1997; 97US-0050934.
PR 06-JUN-1997; 97US-0048970.
PR 05-SEP-1997; 97US-0057765.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ferrie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H,
XX Li H, Li Y, Moore PA, Rosen CA, Ruden SM, Soppet DR;
XX Wei YF, Young PE, Zeng Z;
XX
XX WPI: 1998-520811/44.
XX N-PSDB; AAV34318.
XX
XX Isolated human poly(nucleotide)s encoding secretory peptide(s) -
XX used to develop products for the diagnosis and treatment of e.g.
XX inflammation, cancers, CNS disorders or immune system disorders
XX
XX Claim 1; Page 175; 201pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin Fc
XX portion (e.g. AAV34277) for increasing the stability of the fused
XX protein as compared to the human protein only.
XX The invention relates to 28 novel genes and their fragments (nucleic
XX acid sequences: AAV34286-V34325; amino acid sequences AAW75196-W5235)
XX which are useful for preventing, treating or ameliorating medical
XX conditions e.g. by protein or gene therapy. Also, pathological
XX conditions can be diagnosed by determining the amount of the new
XX polypeptides in a sample or by determining the presence of mutations in
XX the new polynucleotides. Specific uses are described for each of the 28
XX polynucleotides, based on which tissues they are most highly expressed in
XX (see AAV34286 for described uses).
XX
XX Sequence 121 AA:
SQ

```

```

Query Match 54.5%; Score 66; DB 19; Length 121;
Best Local Similarity 100.0%; Pred. No. 2.6e-57;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MRIMLFTAILAFSLAOSGAVCKEPOEEVPGGSKRDPDLYOLLOLFRSHSLEGL 60
DB 1 mrimlftailafslagsfgavckepgeevpggskrdpdllyqlqlfkshslegl 60
OY 61 IKALSO 66
DB 61 lkalsq 66

```

```

RESULT 12
AAV12634
ID AAV12634 standard; Protein: 51 AA.
XX
XX AAV12634;
XX
XX 22-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO: 299 from WO 9905553.
XX
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
KW

```

```

KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; hematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokine; haemostatic;
KW thrombolytic; antiinflammatory; tumour inhibition; antitumour.
XX
XX Homo sapiens.
XX
XX WO9906553-A2.
XX
XX 11-FEB-1999.
XX
XX 31-JUL-1998; 98WO-1B01237.
XX
XX 01-AUG-1997; 97US-0905051.
XX
XX (GEST ) GENSET.
XX
XX Dugclert A, Dumas Milne Edwards J, Lacroix B;
XX
XX WPI: 1999-153783/13.
XX N-PSDB; AAX41492.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
XX cDNA libraries derived from umbilical cord, lymph ganglia,
XX lymphocytes and placental tissue
XX
XX Claim 34; Page 389; 411pp; English.
XX
XX AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAV12521 to
XX AAV12668, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, hematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, antiinflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 51 AA:
SQ

```

```

Query Match 40.5%; Score 49; DB 20; Length 51;
Best Local Similarity 100.0%; Pred. No. 7.6e-41;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MRIMLFTAILAFSLAOSGAVCKEPOEEVPGGSKRDPDLYOLLOL 49
DB 1 mrimlftailafslagsfgavckepgeevpggskrdpdllyqlqlr 49

```

```

RESULT 13
AAW74414
ID AAW74414 standard; Protein: 39 AA.
XX
XX AAW74414;
XX
XX 10-MAY-1999 (first entry)
XX
XX HPMBO91 protein sequence fragment.
XX
XX HPMBO91; neurokinin B precursor; neuronal disorder; CNS related disorder;
XX gastrointestinal disorder; cardiovascular disorder; metabolic disorder;
XX smooth muscle disorder; inflammatory disorder; cancer; gene therapy.
XX
XX Homo sapiens.
OS

```

XX EP892053-A2.
 PN 20-JAN-1999.
 PD
 XX 26-JUN-1998; 98BP-0305066.
 PF
 XX 14-JUL-1997; 97EP-0305215.
 PR
 XX (HUMA-) HUMAN GENOME SCI.
 PA (SMK) SMITHKLINE BEECHAM PLC.
 PI Duckworth DM, Hastings GA, Ruben SM;
 DR MPI: 1999-083570/08.
 XX
 PT New human neurokinin B precursor (HPMBQ91) polypeptides and
 PT polynucleotide - useful as diagnostic reagents and for prevention
 PT and treatment of neuronal, metabolic, inflammatory and
 PT gastrointestinal disorders, and cancers
 XX
 PS Claim 14; Page 15; 18pp; English.
 CC This sequence is the human neurokinin B precursor HPMBQ91 of the
 CC invention. HPMBQ91 polypeptides and polynucleotides are useful for
 CC diagnosing susceptibility to diseases associated with HPMBQ91 protein
 CC imbalance by determining HPMBQ91 polypeptide expression levels. HPMBQ91
 CC polypeptides can be used to screen for agonists and antagonists by
 CC measuring the binding to HPMBQ91, and observing the stimulation or
 CC inhibition of HPMBQ91 function. These can be used in treatment to
 CC activate or inhibit HPMBQ91 activity to treat conditions associated with
 CC a lack of HPMBQ91 protein. Gene therapy may also be used to affect
 CC endogenous HPMBQ91 polypeptide expression. HPMBQ91 antibodies are useful
 CC for inducing an immune response to immunise and prevent disease, and for
 CC isolating HPMBQ91 clones or purifying the polypeptides by affinity
 CC chromatography. HPMBQ91 polypeptides can be administered directly or as a
 CC vaccine to inoculate against disease. Diseases diagnosed, prevented and
 CC treated include: neuronal disorders; CNS related disorders;
 CC gastrointestinal and cardiovascular disorders; metabolic disorders
 CC including diabetes and obesity; smooth muscle disorders; inflammatory
 CC disorders; and cancers including adenomas, leiomyomas, liposarcomas,
 CC melanomas, pulmonary chondroid hamartomas, lung, prostate and breast
 CC cancer. The HPMBQ91 polypeptide is also useful for mapping the gene to a
 CC chromosome, allowing gene inheritance to be studied through linkage
 CC analysis, and tissue localisation studies, for determining HPMBQ91
 CC expression patterns.
 XX
 SQ Sequence 39 AA;
 Query Match 32.2%; Score 39; DB 20; Length 39;
 Best Local Similarity 100.0%; Pred. No. 4,5e-31;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 40 DPDFLQQLQRLFKSHSSLEGALKALSOASTDPKSTSP 78
 DB 1 dpdlyqlqlrllfkshsslegalkalsqastdpkrestpe 39
 RESULT 14
 AAW75249
 ID AAW75249 standard; Protein: 36 AA.
 XX
 AC AAW75249;
 XX
 DT 29-JAN-1999 (first entry)
 DE Fragment of human secreted protein encoded by gene 17.
 XX
 KW Human: secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;

KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN WO9840483-A2.
 XX
 PD 17-SEP-1998.
 XX
 PF 12-MAR-1998; 98WO-US04858.
 XX
 PR 19-DEC-1997; 97US-0068368.
 PR 14-MAR-1997; 97US-0040710.
 PR 14-MAR-1997; 97US-0040762.
 PR 30-MAY-1997; 97US-0048100.
 PR 30-MAY-1997; 97US-0048189.
 PR 30-MAY-1997; 97US-0048357.
 PR 30-MAY-1997; 97US-0050934.
 PR 06-JUN-1997; 97US-0048970.
 PR 05-SEP-1997; 97US-0057765.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ferrite AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;
 PI Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;
 PI Wei YF, Young PE, Zeng Z;
 XX
 DR MPI: 1998-520811/44.
 DR N-PSDB; AAV34302.
 XX
 PT Isolated human poly(nucleotide(s)) encoding secretory peptide(s) -
 PT used to develop products for the diagnosis and treatment of e.g.
 PT inflammation, cancers, CNS disorders or immune system disorders
 XX
 PS Disclosure; Page 19; 201pp; English.
 CC This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule designated Gene 17 (AAV34302). The gene
 CC can be used to generate fusion proteins by linking to the gene to a
 CC human immunoglobulin Fc portion (e.g. AAV34277) for increasing the
 CC stability of the fused protein as compared to the human protein only.
 CC The invention relates to 28 novel genes and their fragments (nucleic
 CC acid sequences: AAV34286-V34325; amino acid sequences AAW75196-W75235)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 28
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV34286 for described uses).
 XX
 SQ Sequence 36 AA;
 Query Match 29.8%; Score 36; DB 19; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3,9e-28;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 77 PEKRDHDFVGLMGKRKSVOPDSEPTDYNQENVSPFG 112
 DB 1 pekrrdmhdfvglmgrksvopdsptdngenvspfsfg 36
 RESULT 15
 AAW75250
 ID AAW75250 standard; Protein: 15 AA.
 XX
 AC AAW75250;
 XX
 DT 29-JAN-1999 (first entry)

Job time: 124 sec

Fragment of human secreted protein encoded by gene 17.

DE Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; fetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

PN W09840483-A2.

PD 17-SEP-1998.

PF 12-MAR-1998; 98WO-US04858.

PR 19-DEC-1997; 97US-0068368.

PR 14-MAR-1997; 97US-0040710.

PR 14-MAR-1997; 97US-0040762.

PR 30-MAY-1997; 97US-0048100.

PR 30-MAY-1997; 97US-0048189.

PR 30-MAY-1997; 97US-0048357.

PR 06-JUN-1997; 97US-0050934.

PR 05-SEP-1997; 97US-0048970.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ferlie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;

PI Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;

PI Wei YF, Young PE, Zeng Z;

DR MPI, 1998-520811/44.

DR N-PSDB; AAV34302.

PT Isolated human poly:nucleotide(s) encoding secretory peptide(s)

PT used to develop products for the diagnosis and treatment of e.g.

PT inflammation, cancers, CNS disorders or immune system disorders

PS Disclosure; Page 19; 201pp; English.

XX This sequence represents a fragment of a secreted human protein encoded

CC by the nucleic acid molecule designated Gene 17 (AAV34302). The gene

CC can be used to generate fusion proteins by linking to the gene to a

CC human immunoglobulin Fc portion (e.g. AAV34277) for increasing the

CC stability of the fused protein as compared to the human protein only.

CC The invention relates to 28 novel genes and their fragments (nucleic

CC acid sequences: AAV34286-V34325; amino acid sequences AAW5196-W75235)

CC which are useful for preventing, treating or ameliorating medical

CC conditions e.g. by protein or gene therapy. Also, pathological

CC polypeptides in a sample or by determining the amount of the new

CC the new polynucleotides. Specific uses are described for each of the 28

CC polynucleotides, based on which tissues they are most highly expressed in

CC (see AAV34286 for described uses).

XX Sequence 15 AA;

Query Match 12.4%; Score 15; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 79 KRDMDPFYGLMGKR 93
 |||||
 Db 1 Krdmddfvgimgkr 15

Search completed: May 24, 2002, 17:06:10

